MATERIAL DIDÁCTICO
MATEMÁTICAS

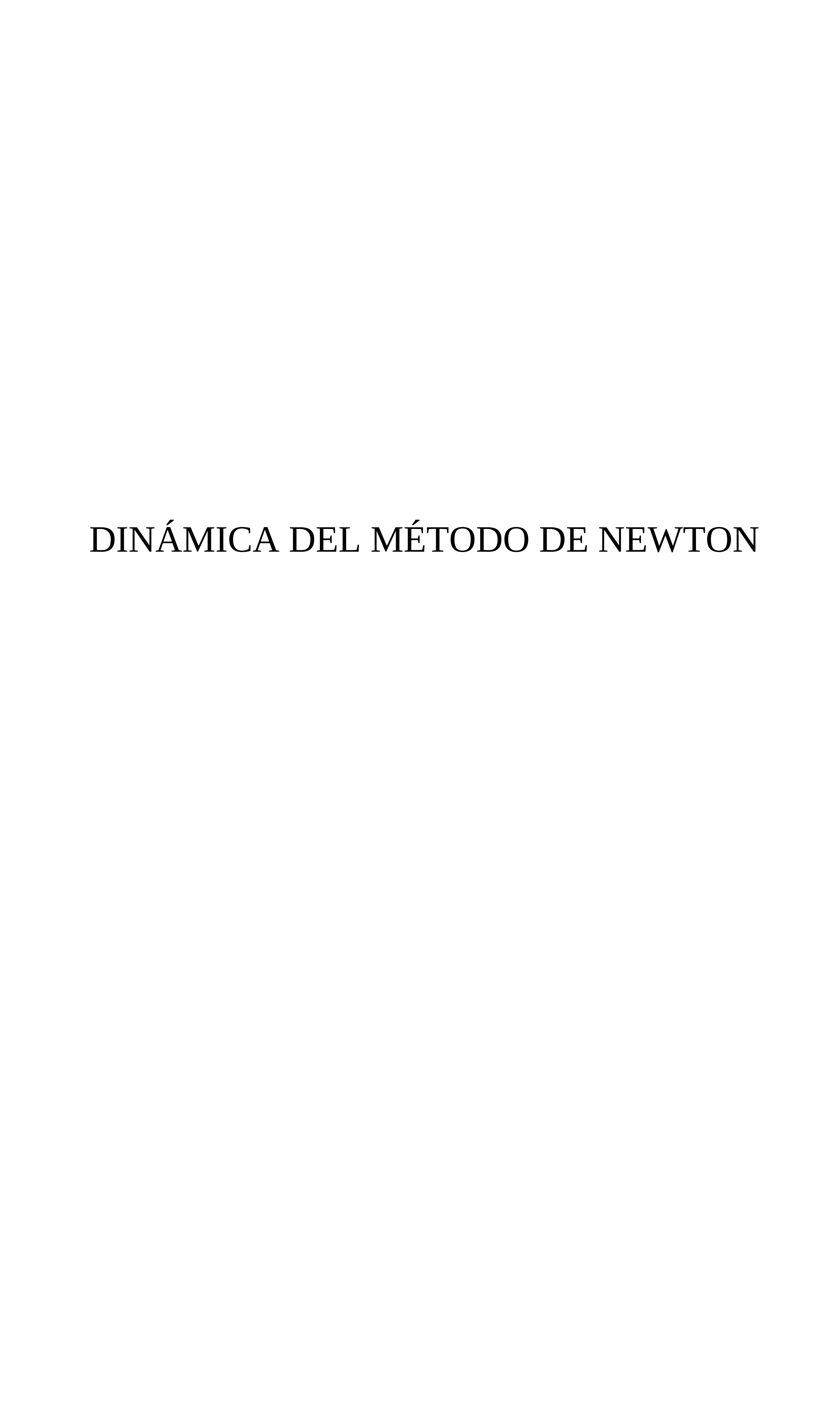
9

DINÁMICA DEL MÉTODO DE NEWTON

Sergio Plaza Salinas
José Manuel Gutiérrez Jiménez







DINÁMICA DEL MÉTODO DE NEWTON

MATERIAL DIDÁCTICO

Matemáticas

nº 9

Sergio Plaza Salinas

José Manuel Gutiérrez Jiménez

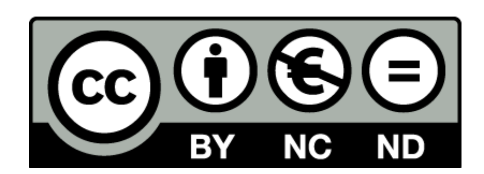
DINÁMICA DEL MÉTODO DE NEWTON

UNIVERSIDAD DE LA RIOJA

SERVICIO DE PUBLICACIONES

2013

PLAZA SALINAS, Sergio
Dinámica del método de Newton [Recurso electrónico] / Sergio Plaza Salinas y José Manuel Gutiérrez Jiménez. – Logroño : Universidad de La Rioja, Servicio de Publicaciones, 2013.
XII, 210 p. ; v. digital. – (Material didáctico. Matemáticas ; 09)
ISBN 978-84-695-7461-4
1. Sistemas dinámicos diferenciales. 2. Métodos iterativos. I. Gutiérrez Jiménez, José Manuel. II. Título. III. Universidad de La Rioja. Servicio de Publicaciones. IV. Serie.
517.9
PBKS – IBIC 1.1



publicaciones.unirioja.es

E-mail: publicaciones@unirioja.es

ISBN 978-84-695-7461-4

Edita: Universidad de La Rioja, Servicio de Publicaciones

Dedicatoria



Figura 1: Los autores fotografiados en su visita a la casa de Pablo Neruda en Isla Negra, en abril de 2011.

Este libro tuvo su origen en un viaje en tren entre Cartagena y Madrid, cuando los dos autores volvíamos de una tesis doctoral en la ciudad de origen, allá por marzo de 2009. Mientras disfrutábamos contemplando los paisajes que en su día viera Don Quijote y mientras hablábamos de lo humano y lo divino, se nos ocurrió escribir este texto. Nuestra idea inicial fue mostrar la estrecha relación que existe entre dos áreas muy activas en investigación, como lo son los métodos iterativos para aproximar raíces de ecuaciones no lineales y los sistemas dinámicos, en este caso discretos. Inicialmente, pensamos en incluir varios métodos, en especial aquéllos más populares, pero con el avance del escrito, nos fuimos restringiendo al método de Newton.

No nos resultó complicado trabajar en la distancia y, poco a poco, el libró empezó a ir tomando forma. Incluso, vimos que iba creciendo mucho. Así que, para no extender demasiado el texto, se ha quedado mucho material sin tratar. Además tuvimos la suerte de poder trabajar mano a mano, con dos visitas de Sergio a Logroño y una visita mía a Santiago de Chile, en abril de 2011, donde el libro quedó prácticamente terminado. Poco hacía sospechar que unos meses después, en junio de 2011, una terrible enfermedad apartara a Sergio del camino de la vida.

A partir de entonces, después de disipar las dudas sobre si continuar o no con el proceso, decidí que la mejor manera de hacer justicia con Sergio era publicar este texto que él había gestado y en el que tenía puestas tantas ilusiones. Dos años más tarde, el libro está listo para su publicación. Sería el momento ahora de hacer los agradecimientos a las personas e instituciones que colaboraron e hicieron posible llevar a cabo esta empresa: familiares, amigos, colegas, revisores, fuentes de financiación, etc., pero espero que todos ellos entiendan que los agradecimientos y la dedicatoria de este libro vayan para Sergio Plaza Salinas.

¡Va por ti, Sergio!

José Manuel Gutiérrez

Logroño, 8 de mayo de 2013

Prólogo

Lo primero que es obligatorio hacer, es alertar al lector que el título de esta obra en sí, conlleva la conjunción de dos grandes vertientes de las matemáticas modernas. A saber, los sistemas dinámicos y los métodos numéricos para encontrar raíces de ecuaciones. Ambas, ramas de la matemática de fecundo desarrollo en el siglo XX e inicio de este nuevo siglo XXI.

Así, la sola escritura de un libro como este, no resulta ser una tarea fácil. Esto, porque cada una de estas ramas tiene sus particularidades. Los sistemas dinámicos, por así decirlo, son la matemática de la intuición, por excelencia, y los métodos numéricos vienen del análisis, por origen. Así, introducirse a este tema requiere de las dos capacidades que se han enfrentado desde siglos en matemática: la intuición y el pensamiento analítico.

Pero, no es tan verdad que deban siempre enfrentarse. Y este libro aquí es un ejemplo de ello. Pueden, también, complementarse, como se complementaron por años José Manuel y Sergio en su trabajo original como matemáticos. Así invito al virtual lector, a introducirse en esta obra para que conozca como se hace esto de mezclar los sistemas dinámicos con los métodos numéricos.

El primer capítulo es una breve introducción a los sistemas dinámicos. Al lector no familiarizado le recomiendo rehacer las cuentas, sobre todo lo relacionado con la familia cuadrática, porque con ello empezará a sentir el «sabor» del tipo de problema que se aborda en dinámica.

El segundo capítulo introduce el método de Newton, su historia, entrega una idea de sus generalizaciones y se concentra en él, su convergencia y diversas aplicaciones. Al leerlo, uno se da cuenta de cuál es el tipo de problema que se quiere abordar con estas técnicas, como se hace y como se aplican.

Ya el capítulo tres trata una introducción el tema principal del libro: la dinámica del método de Newton y lo hace en el primer lugar donde se debe ver, la recta real. Así, se estudia la dinámica de la transformación de Newton en al ámbito del conjunto de los números reales y se completa el capítulo con las bifurcaciones del referido método. Para este capítulo ya se hace necesario algún mejor dominio de la matemática. Digamos: un curso bien hecho de introducción al análisis real; otro de álgebra intermedia, algo de polinomios, álgebra lineal y un poco de topología.

El capítulo cuatro, que trata sobre la dinámica del método de Newton en el campo de los números complejos, es la parte más compleja del libro y tiene que ver con: una breve

introducción a la dinámica compleja; el estudio del método de Newton aplicado a polinomios de grado 2, 3, 4 y 5 y la determinación de algoritmos generalmente convergentes para polinomios complejos. Esta última parte tiene que ver con desarrollos relativamente recientes y requerirá mayor concentración del lector. En todo caso, hay una rica bibliografía que puede ayudar a comprender de mejor manera esta última parte.

El libro concluye con un «cogollo» estético sobre comportamiento dinámico de polinomios y sus subyacentes conjuntos de Julia, de Mandelbrot y la sorprendente relación que dice que conjuntos de Julia llenos de polinomios cuadráticos, aparecen como parte del conjunto de Fatou del método de Newton aplicados a polinomios cúbicos. Concluye la obra mostrando el fractal de Chicho (Chicho es el sobrenombre del matemático español José Javier Guadalupe (1945–2000)), que puede obtenerse jugando con una modificación de la función de iteración asociada al conjunto de Mandelbrot del polinomio cuadrático $z^2 + c$.

En fin, damos la bienvenida a este libro que resulta ser un lugar para aprender y entretenerse, para quien quiera hacerlo.

Dr. Rafael Labarca B.
Profesor de Matemática
Universidad de Santiago de Chile

Santiago de Chile, 6 de mayo de 2013

Índice general

Capítulo 1

Sistemas dinámicos discretos

1.1. Introducción

Una parte de la teoría de los sistemas dinámicos se dedica a estudiar el comportamiento de las iteraciones sucesivas de una función en términos de los diferentes puntos iniciales.

El problema de determinar la conducta de los iterados de una función aparece frecuentemente en algunos estudios naturales, como por ejemplo en la dinámica de poblaciones. El objetivo de este tipo de problemas es determinar el número de individuos $x(t)$ de una determinada población como función del tiempo t. Una aproximación a este problema, debida al economista inglés Thomas R. Malthus («Primer ensayo de la población», 1798), consiste en asumir que existe una tasa de nacimiento constante b, lo cual significa que en una población de x individuos podemos esperar que bx nuevos individuos nazcan en el transcurso de una unidad de tiempo. Esto puede ser formulado como una ecuación en diferencias finitas

$$x_{n+1} = x_n + bx_n = (1 + b)x_n,$$

donde x_n es la población en el tiempo n. Iterando esta relación obtenemos la expresión

$$x_n = (1 + b)^n x_0$$

como predicción de la población, donde x_0 es la población inicial.

Frecuentemente, se supone que la población sigue una evolución continua, de modo que idealmente las mediciones x_n serán tomadas en intervalos pequeños de longitud h. Esto lleva a la ecuación de diferencias

$$x((n + 1)h) - x(nh) = bhx(nh).$$

Escribiendo $x(t)$ como función de $t = nh$, y reordenando, obtenemos

$$x'(t) = \lim_{h \to 0} \frac{x(t + h) - x(t)}{h} = bx(t).$$

Como vemos, la ecuación en diferencias anterior se aproxima a la ecuación diferencial

$$x'(t) = bx(t),$$

conocida como *ecuación diferencial de crecimiento exponencial.* La solución de esta ecuación es la función exponencial

$$x(t) = x_0 e^{bt}.$$

En este caso, las versiones discreta y continua de este sistema dinámico se comportan más o menos de la misma forma.

Nótese que en el modelo anterior no hemos considerado la tasa de mortalidad. Podemos asumir que ésta es proporcional a la población existente en cada momento dado. Esto nos lleva otra vez a la ecuación de diferencias de crecimiento exponencial si la tasa de nacimientos excede a la tasa de muertes, o al decaimiento exponencial si ocurre lo opuesto. Sin embargo, uno puede suponer que la tasa de muerte es proporcional al número de encuentros entre individuos. Esto es, cuanto más y más individuos entren en contacto unos con otros, un mayor porcentaje de individuos muere. Esta es la base del conocido como *modelo de Verhulst*, que en su versión continua da lugar a la ecuación diferencial

$$x'(t) = bx(t) - dx^2(t),$$

también conocida como *ecuación logística*, y cuya solución es

$$x(t) = \frac{bx_0}{dx_0 + (b - dx_0)e^{-bt}}.$$

Obsérvese que las muertes decaen exponencialmente a 0 cuando $t \to \infty$. Por lo tanto, este modelo predice una población límite de b/d individuos, independiente de la población inicial.

La expresión, relativamente sencilla, de la función logística no hace sospechar que esconda una riqueza dinámica extraordinaria. Por ejemplo, una variante de la función logística en su versión discreta,

$$x_{n+1} = (1 + b)x_n - dx_n^2$$

fue empleada por el biólogo Robert May en 1976 para estudiar el crecimiento de una población de insectos [97]. May constató que el modelo así obtenido podía presentar comportamientos muy variados, como vemos en el siguiente ejemplo.

Ejemplo 1.1. *Analícese el comportamiento de la sucesión*

$$x_{n+1} = cx_n(1 - x_n)$$

para distintos valores del parámetro c, considerando en todos los casos el mismo punto de partida: $x_0 = 0.1$.

Para $c = 0.75$ la sucesión $\{x_n\}$ converge a 0. Para $c = 2.5$ la sucesión $\{x_n\}$ converge a 0.6. Para $c = 3.2$, y a partir de un cierto valor de n, se observa que la sucesión $\{x_n\}$ va tomando los valores 0.513045 y 0.799455. Sin embargo, para $c = 3.5$ esta tendencia repetitiva ha aumentado a cuatro términos: 0.38280, 0.826941, 0.500884 y 0.874997. Por último para $c = 3.8$ los términos de la sucesión anterior no siguen ningún orden aparente. Presentan, por tanto, un comportamiento caótico.

1.2. Conceptos básicos

En esta sección presentamos algunos de los conceptos básicos de la dinámica de funciones escalares. Sólo consideramos los resultados que, a nuestro juicio nos han parecido más relevantes o que vayamos a utilizar en secciones posteriores. Existe una extensa bibliografía que permite profundizar más en este apartado. A modo de ejemplo, se pueden consultar los textos [3, 43, 57, 66, 69, 94, 127, 126, 141] y algunos artículos [91, 97, 144].

La primera definición hace referencia a lo que se entiende por un *sistema dinámico discreto*.

Definición 1.1. *Un* sistema dinámico discreto *es un par* (X, f) *formado por un espacio métrico* X *y una función* $f : X \to X$.

En esta sección nos centramos en el caso de que el espacio métrico X es el conjunto $\mathbb{R}$ de los números reales. No obstante, muchas de las definiciones y conceptos que vamos a introducir aquí se pueden extender de forma natural a otros espacios como los números complejos $\mathbb{C}$ o a los espacios $\mathbb{R}^n$ con $n > 1$. Del mismo modo, en un principio, sólo requerimos que las funciones sean continuas, aunque también se pueden obtener resultados específicos cuando se trabaja con funciones derivables.

Sea $f : I \to \mathbb{R}$ una función continua definida en un intervalo $I \subset \mathbb{R}$. En lo que sigue f^n denotará la composición de f con sí misma n veces, es decir,

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ veces}}$$

cuando esta composición tenga sentido.

Definición 1.2. *Sea* $p \in I$. *Si las sucesivas imágenes de* p, $f(p), f^2(p), f^3(p), \ldots$, *están definidas, se define la* órbita de p *como el conjunto*

$$\text{orb}(p) = \{p, f(p), f^2(p), \ldots, f^n(p), \ldots\}.$$

La teoría de los sistemas dinámicos trata de analizar el comportamiento de $\text{orb}(p)$ para diferentes valores de $p \in I$. En una primera clasificación de las órbitas podemos distinguir aquéllas que se acercan a un cierto valor o las que presentan un comportamiento periódico. Desde el punto de vista dinámico, tal vez, las más interesantes resulten las órbitas no periódicas y que puedan dar lugar a un comportamiento caótico.

Comenzamos definiendo los elementos más simples en el estudio de los sistemas dinámicos discretos: los puntos fijos y los puntos periódicos.

Definición 1.3. *Sea $p \in I$. Decimos que*

(a) p es un punto fijo de f *si $f(p) = p$.*

(b) p es un punto periódico de período k *si, $f^k(p) = p$ y $f^j(p) \neq p$ para $1 \leq j \leq k-1$.*

(c) Nos referiremos a la órbita periódica de un punto periódico de período k como un k-ciclo.

A continuación hacemos algunas observaciones interesantes acerca de esta definición.

(i) Los puntos fijos de f son aquéllos en donde su gráfico intersecta a la línea $y = x$.

(ii) Los puntos periódicos de período $k \geq 2$ de f son los puntos fijos de f^k que no son puntos fijos de f^j para $1 \leq j \leq k-1$.

(iii) Si p es un punto periódico de f entonces su órbita es finita. En este caso, el número de elementos en $\text{orb}(p)$ es su período.

Ejemplo 1.2. *Para $f : \mathbb{R} \to \mathbb{R}$, $f(x) = -x^3$, $x_0 = 0$ es un punto fijo y $x_1 = 1$ es un punto periódico de período 2.*

Aunque existe una gran variedad de resultados que garantizan la existencia de puntos fijos de una función bajo distintas suposiciones, los siguientes teoremas, son quizás los más elementales sobre existencia de puntos fijos. Omitimos su demostración, que es una consecuencia inmediata del teorema del valor intermedio (véase [80]).

Teorema 1.1. *Sean $I = [a, b]$ un intervalo cerrado y $f : I \to \mathbb{R}$ una aplicación continua tal que $f(I) \supseteq I$ ó $I \supseteq f(I)$. Entonces f tiene un punto fijo en I.*

Ejemplo 1.3. *La función $f(x) = 1 - x^2$ definida en el intervalo $I = [0, 1]$ satisface las condiciones del teorema 1.1.*

> En efecto, notemos que $f(I) = I$, luego f tiene un punto fijo en I. En este caso, dicho punto fijo es fácil de encontrar. Para ello tenemos que resolver la ecuación $f(x) = x$, es decir, $1 - x^2 = x$, de donde $x_0 = (-1 + \sqrt{5})/2$. Notemos que existe otra solución de la ecuación anterior, pero ésta no pertenece al intervalo $[0, 1]$.

Definición 1.4. *Sean $p \in I$ y $f : I \to \mathbb{R}$. Decimos que*

(a) p es eventualmente fijo *si existe N tal que $f^{n+1}(p) = f^n(p)$ para $n \geq N$.*

(b) p es eventualmente periódico con período k *si existe N tal que $f^{n+k}(p) = f^n(p)$ para $n \geq N$. En otras palabras, $f^N(p)$ está sobre una órbita periódica.*

Notemos que p es un punto eventualmente fijo de f si es una preimagen de un punto fijo. En otras palabras, existe un punto fijo q de f tal que $p \in \bigcup_{n\geq 1} f^{-n}(q)$, donde

$$f^{-k}(q) = \{x \in I : f^k(x) = q\}.$$

Ejemplo 1.4. *La función $f(x) = |1-x|$, definida para $x \in \mathbb{R}$, tiene un punto eventualmente fijo en $p = 3/2$ y un punto eventualmente periódico de período 2 en $p = 2$.*

En efecto, como $f(1/2) = 1/2$ y $f(3/2) = 1/2$, se tiene que $f^{n+1}(3/2) = f^n(3/2) = 1/2$ si $n \geq 1$. Por otra parte, como $\{0, 1\}$ es una órbita periódica de período 2, y como $f(2) = 1$ se sigue que 2 es eventualmente periódico de período 2.

Ejemplo 1.5. *La función $f(x) = x^3$ tiene un punto fijo en $x = 0$. Además, los puntos en $x \in (-1, 1)$ satisfacen $\lim_{n\to\infty} f^n(x) = 0 = f(0)$ (véase la figura 1.1).*

Por otra parte, la función $f(x) = \sqrt[3]{x}$ también tiene un punto fijo en $x = 0$. Sin embargo, la órbita de cualquier valor $x_0 \neq 0$, se aleja de 0, incluso aunque x_0 esté muy próximo a 0 (véase la figura 1.1).

Para realizar un análisis gráfico de la órbita de un punto x_0 por una función de iteración $f(x)$, se puede proceder de la siguiente manera:

1. Representar con los mismos ejes el gráfico de la función $y = f(x)$ y el de la diagonal $y = x$.
2. Situar el punto x_0 en ele eje de abcisas y desplazarse verticalmente hasta cortar con la gráfica de $y = f(x)$, obteniendo un punto cuya abcisa es $f(x_0)$.
3. A partir de ese punto, desplazarse horizontalmente hasta intersecar la diagonal $y = x$.
4. Volver a desplazarse verticalmente hasta encontrar de nuevo la gráfica de la curva $y = f(x)$. Obtenemos ahora un punto cuya abcisa es $f(f(x_0))$.
5. Repetir el proceso las veces que se consideren oportunas.

El ejemplo anterior motiva una clasificación de los puntos fijos atendiendo a su carácter atractor o repulsor respecto a las órbitas de los puntos cercanos.

Definición 1.5. *Sea p un punto fijo de f. Decimos que p es un* atractor *si existe un intervalo abierto $J \ni p$ tal que para cada $x \in J$, $\lim_{n\to\infty} f^n(x) = p$.*

Por el contrario decimos que p es un repulsor *si existe un intervalo abierto $J \ni p$ tal que para cada $x \in J$, $x \neq p$, existe $k \in \mathbb{N}$ tal que $f^k(x) \notin J$.*

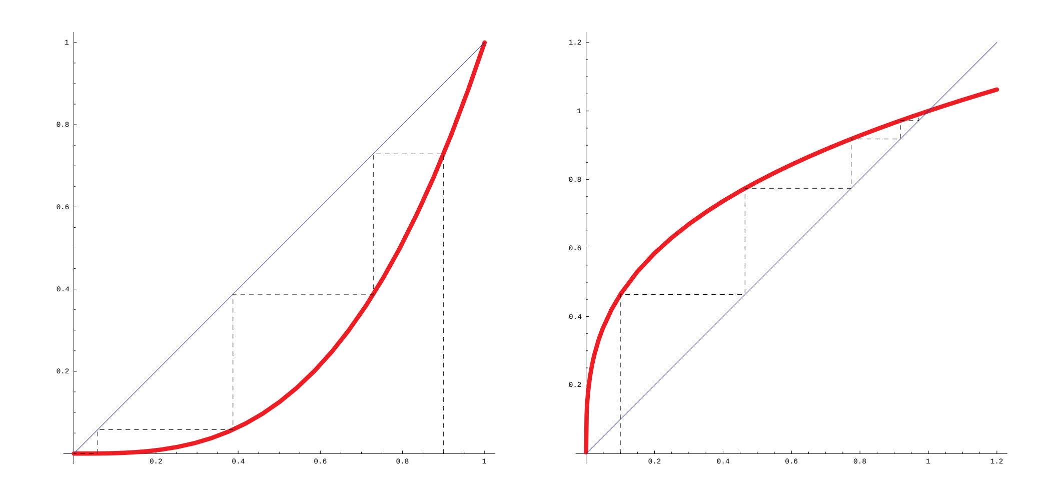


Figura 1.1: Análisis gráfico del comportamiento atractor de $p = 0$ como punto fijo de $f(x) = x^3$ y del comportamiento repulsor de $p = 0$ como punto fijo de $f(x) = x^{1/3}$. En el primer caso se muestra cómo la órbita del punto $x_0 = 0.9$ converge hacia 0, mientras que en el segundo caso se muestra cómo la órbita del punto $x_0 = 0.1$ se «escapa» del 0, siendo atraída por otro punto fijo, en este caso $p = 1$.

Estas definiciones se pueden extender a los puntos periódicos de f. En concreto, p es un *punto periódico atractor (repulsor)* de f si y sólo si p es un punto fijo atractor (repulsor) de f^k, siendo k el período de p.

Nótese que hay puntos fijos que no son ni atractores ni repulsores. Por ejemplo $p = 0$ es un punto fijo de $f(x) = x + x^2$. Si $x_0 \in [-1, 0]$ la órbita de x_0 converge a 0. Si $x_0 \notin [-1, 0]$, su órbita es divergente a ∞.

Definición 1.6. *Sea p un punto fijo de f. La* cuenca de atracción de p *es el conjunto* $\mathcal{B}(p) = \{x : f^n(x) \to p, \text{ cuando } n \to \infty\}$.

La cuenca de atracción inmediata de p *es la componente conexa de $\mathcal{B}(p)$ que contiene a p.*

Definición 1.7. *Sea p un punto periódico de período k de f. La* cuenca de atracción de p *es* $\mathcal{B}(p) = \{x : f^{nk}(x) \to p, \text{ cuando } n \to \infty\}$.

La cuenca de atracción de la órbita de p, *es* $\mathcal{B}(\mathrm{orb}(p)) = \bigcup_{j=0}^{k-1} f^j(\mathcal{B}(p)) = \bigcup_{j=0}^{k-1} \mathcal{B}(f^j(p))$.

La cuenca de atracción inmediata de la órbita de p *es la unión de las componentes conexas que contienen los puntos de la órbita de p.*

Es claro que si p es un punto periódico atractor de f, entonces tanto $\mathcal{B}(p)$ como $\mathcal{B}(\mathrm{orb}(p))$ son conjuntos abiertos. Esto no ocurre para puntos fijos no atractores.

Por ejemplo, para $f(x) = -x^3$, el punto $p = 0$ es un punto fijo atractor, y $\mathcal{B}(0) = (-1, 1)$. Por otra parte, para $f(x) = x + x^2$, el punto $p = 0$ es un punto fijo no atractor, y $\mathcal{B}(0) = [-1, 0]$, el cual no es abierto.

Teorema 1.2. *Las cuencas de atracción de puntos periódicos distintos no se intersecan.*

Quizás los resultados más sorprendentes en dinámica uno dimensional, con la sola hipótesis de continuidad son los siguientes (véase [43], [91]).

Teorema 1.3 (Li y Yorke). *Si una función continua $f : I \to \mathbb{R}$ tiene un punto periódico de período tres, entonces f tiene puntos periódicos de todos los períodos.*

Ejemplo 1.6. *Considérese la función $f(x) = (-3x^2 + 5x + 2)/2$. Como $f(0) = 1$, $f(1) = 2$ y $f(2) = 0$, se tiene que $\{0, 1, 2\}$ forma una órbita periódica de período 3. En consecuencia, por el teorema de Li y Yorke, f tiene puntos periódicos de todos los períodos.*

Definición 1.8. *El* orden de Sarkovskii *de los números naturales es:*

$$3 \succ 5 \succ 7 \succ 9 \succ \cdots \succ 2^0 \cdot (2n+1) \succ \cdots$$

$$\cdots \succ 2 \cdot 3 \succ 2 \cdot 5 \succ 2 \cdot 7 \succ 2 \cdot 9 \succ \cdots \succ 2^1 \cdot (2n+1) \succ \cdots$$

$$\cdots \succ 2^2 \cdot 3 \succ 2^2 \cdot 5 \succ 2^2 \cdot 7 \succ 2^2 \cdot 9 \succ \cdots \succ 2^2 \cdot (2n+1) \succ \cdots$$

$$\cdots \succ 2^{n+1} \succ 2^n \succ \cdots \succ 2^3 \succ 2^2 \succ 2 \succ 1,$$

donde $n \succ m$ significa que n precede a m en el orden.

Teorema 1.4 (Sarkovskii). *Supongamos que $f : \mathbb{R} \to \mathbb{R}$ es continua y que f tiene un punto periódico con período n. Si $n \succ m$ en el orden de Sarkovskii, entonces se prueba que f tiene un punto periódico con período m.*

Ahora, en lo que sigue asumiremos que las funciones en cuestión son al menos diferenciables. En este caso, se pueden dar caracterizaciones de los puntos fijos de una función en términos de su derivada.

Teorema 1.5. *Sea I un intervalo cerrado. Si $f : I \to I$ es diferenciable y satisface $|f'(x)| < 1$ para todo $x \in I$, entonces f tiene un único punto fijo en I.*

Demostración. Inmediata a partir del teorema del valor medio. □

Supongamos que f es de clase $C^1(I)$, es decir, derivable en I y con la función derivada continua. Sea p un punto fijo de f contenido en el intervalo I. El carácter de la derivada de f en p nos proporciona información sobre el punto fijo. Notemos que si $|f'(p)| < 1$, entonces existe un intervalo abierto $U \subseteq I$ con $p \in U$, tal que para cada $x \in U$ se tiene $\lim_{n\to\infty} f^n(x) = p$. En consecuencia $U \subset \mathcal{B}(p)$.

Sin embargo, si $|f'(p)| > 1$ existe un intervalo abierto conteniendo a p tal que todos los puntos de ese intervalo, diferentes del punto p, salen del intervalo bajo iteraciones por f.

Estas observaciones están en la base del siguiente resultado que caracteriza los puntos fijos de una función en términos de su derivada.

Teorema 1.6. *Sea $f : I \to \mathbb{R}$ una función de clase $C^1(I)$ y sea p un punto fijo de f.*

1. *Si $|f'(p)| < 1$, entonces p es un punto fijo atractor de f.*
2. *Si $|f'(p)| > 1$, entonces p es un punto fijo repulsor.*

Definición 1.9. *Sea f una función derivable en un entorno de un punto fijo p. Decimos que p es un* punto fijo indiferente (o neutro) *de f si $|f'(p)| = 1$.*

Además decimos que p es un punto periódico indiferente (o neutro) *de f si p es un punto fijo indiferente (o neutro) de f^k, siendo k el período de p.*

Notemos que cuando p es un punto fijo atractor, el valor $|f'(p)|$ proporciona información sobre la velocidad de convergencia de las órbitas $\text{orb}(x)$ para puntos x cercanos a p. Cuanto menor sea $|f'(p)|$, más rápida será dicha convergencia. En concreto, si $f'(p) = 0$, se dice que p es un *punto fijo superatractor*.

Ejemplo 1.7. *Estudiemos el carácter de $x = 0$ como punto fijo de las siguientes funciones: $f(x) = x^3$, $f(x) = \sqrt[3]{x}$, $f(x) = x - x^3$, $f(x) = x + x^3$ y $f(x) = x + x^2$.*

En el primer caso, $f'(0) = 0$, luego $x = 0$ es un punto fijo superatractor.

Sin embargo, para la función $f(x) = \sqrt[3]{x}$, $\lim_{x\to 0} f'(x) = \infty$, luego $x = 0$ es un punto fijo repulsor.

Es inmediato comprobar que para las funciones $f(x) = x - x^3$, $f(x) = x + x^3$ y $f(x) = x + x^2$ se tiene $f'(0) = 1$, luego se trata de un punto fijo indiferente. Sin embargo, en el primer caso, el punto fijo tiene un carácter atractor, en el segundo caso el punto fijo es repulsor y en el tercer caso el punto fijo tiene un carácter atractor para las órbitas de puntos $x \in [-1, 0]$, mientras que el carácter es repulsor para las órbitas de puntos $x > 0$.

Como vemos en el caso de los puntos fijos indiferentes, el comportamiento puede ser de lo más variado. Se podría obtener más información sobre el carácter del punto fijo analizando el crecimiento y decrecimiento de la función $f'(x)$. En concreto, si p es un punto fijo indiferente de $f(x)$ y la función $|f'(x)|$ presenta un máximo local (mínimo local) en $x = p$, entonces p es un punto fijo atractor (repulsor).

Las funciones con una dinámica más sencilla son las lineales. Los siguientes ejemplos nos muestran su comportamiento dinámico.

Ejemplo 1.8. *Estudio del comportamiento dinámico de las funciones $f(x) = ax$, en función del parámetro $a \in \mathbb{R}$.*

Para cada $x \in \mathbb{R}$ denotamos $a_n(x) = f^n(x) = a^n x$. Podemos distinguir los siguientes casos:

- Si $|a| < 1$, $a_n(x) \to 0$, para todo $x \in \mathbb{R}$. En este caso, sólo hay un punto fijo, $p = 0$, que es atractor.
- Si $|a| > 1$, $|a_n(x)| \to \infty$ para todo $x \in \mathbb{R} - \{0\}$. En este caso, sólo hay un punto fijo, $p = 0$, que es repulsor.
- Si $a = 1$, $a_n(x) = x$ para todo $x \in \mathbb{R}$. En este caso, todos los puntos son fijos y cada uno de ellos es indiferente..
- Si $a = -1$, $a_n(x) = (-1)^n x$ para todo $x \in \mathbb{R}$. En este caso, cada punto distinto de cero está dentro de un 2-ciclo de la forma $\{x, -x\}$.

Ejemplo 1.9. *Estudio del comportamiento dinámico de las funciones de la forma* $f(x) = ax + b$*, con* $a, b \in \mathbb{R}$*,* $b \neq 0$*.*

- Si $a = 1$, no hay ningún punto fijo, las órbitas de un punto x son de la forma $x + nb$. Por lo tanto, si $b > 0$ estas órbitas divergen a ∞ y si $b < 0$ estas órbitas divergen a $-\infty$.
- Si $a \neq 1$, existe un único punto fijo: $p = b/(1 - a)$. El carácter de este punto fijo depende de a:
  - Si $|a| < 1$ es un punto fijo atractor.
  - Si $|a| > 1$ es un punto fijo repulsor.
  - Si $a = -1$, $p = b/2$ es el único punto fijo. El resto de los puntos describen 2-ciclos de la forma $\{x, b - x\}$.

El estudio de la dinámica de una función de la forma $f(x) = ax + b$, con $|a| < 1$ tiene una interesante aplicación en el cálculo de las sumas de series geométricas [143]. En efecto, como se puede ver en la primera gráfica de la figura 1.2, las sumas parciales de la serie

$$b + br + br^2 + \cdots + br^n + \cdots$$

son precisamente los distintos términos de la órbita del cero por las función $f(x) = rx + b$, con $0 < r < 1$. Por lo tanto, se puede probar así que las sumas de la serie geométrica anterior es el punto fijo de $f(x)$, es decir, $b/(1 - r)$.

De forma parecida, tal y como se muestra en la segunda gráfica de la figura 1.2, la suma de la serie geométrica alternada

$$b - br + br^2 + \cdots + b(-r)^n + \cdots$$

es el punto fijo de la función $f(x) = -rx + b$, con $0 < r < 1$, es decir, $b/(1 + r)$.

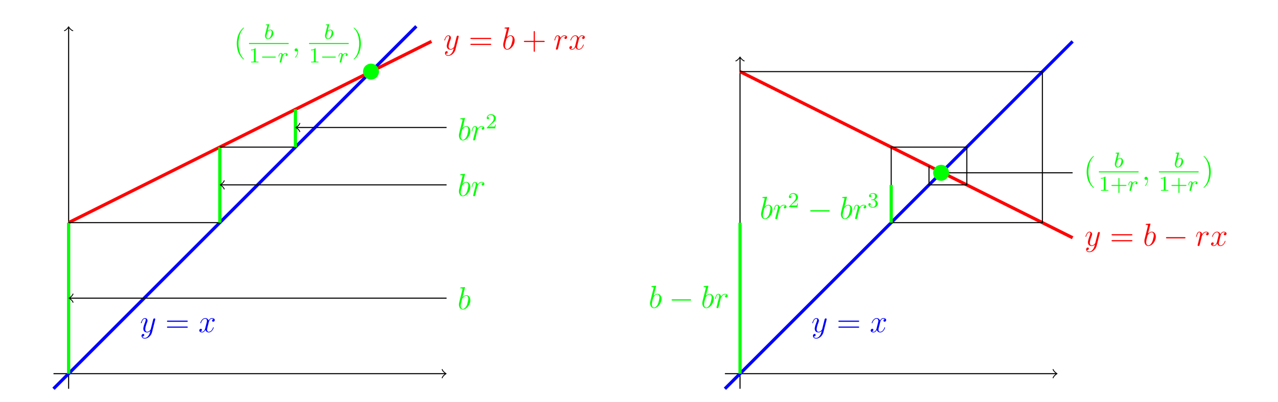


Figura 1.2: Sumas de una serie geométrica y de una serie geométrica alternada como iteraciones de punto fijo.

1.3. Conjugación Topológica

Uno de los conceptos básicos en el estudio de la dinámica de una función es el de *conjugación topológica*, el cual damos a continuación. En líneas generales, la conjugación topológica permite reducir el estudio dinámico de algunas familias de funciones a algunas situaciones concretas.

Definición 1.10. *Sean $f : D \to D$ y $g : E \to E$ dos funciones. Decimos que ellas son* topológicamente conjugadas *si existe un homeomorfismo*[1] $\varphi : D \to E$ *tal que* $\varphi \circ f = g \circ \varphi$, *es decir*

$$\varphi(f(z)) = g(\varphi(z)), \quad z \in D.$$

En este caso, φ *se llama* conjugación topológica *entre* f *y* g.

La definición anterior se traduce en el diagrama conmutativo que se muestra en la figura 1.3. Sin pretender extendernos mucho en este apartado, se puede resaltar el hecho de que la conjugación topológica entre dos funciones envía los puntos fijos de una a los de la otra, los puntos periódicos a los puntos periódicos, los puntos con un comportamiento asintótico a puntos con el mismo comportamiento y, en general, la órbita de un punto $x_0 \in D$ en la órbita de su transformado $\varphi(x_0) \in E$.

El siguiente teorema resume los hechos básicos de la relación entre la dinámica de dos funciones topológicamente conjugadas. Mostramos aquí unas líneas generales de su demostración.

Teorema 1.7. *Sean $f : D \to D$ y $g : E \to E$ dos funciones, y sea $\varphi : D \to E$ una conjugación topológica entre f y g. Entonces*

(a) $\varphi^{-1} : E \to D$ es también una conjugación topológica entre g y f.

[1] Aplicación continua con inversa continua.

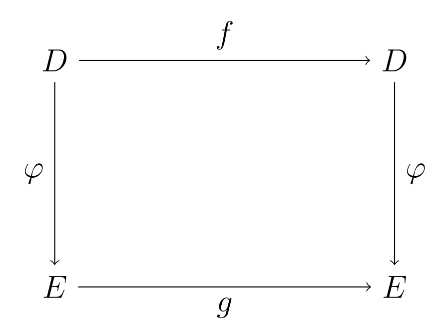


Figura 1.3: Diagrama resultante de una conjugación topológica.

(b) $\varphi \circ f^n = g^n \circ \varphi$ *para todo* $n \in \mathbb{N}$.

(c) p *es un punto periódico de* f *si y sólo si* $\varphi(p)$ *es un punto periódico de* g. *Además,* p *y* $\varphi(p)$ *tienen períodos iguales.*

(c') Para funciones derivables, si p *es un punto periódico de* f *y* φ' *no se anula en la órbita de* p, *entonces* p *y* $\varphi(p)$ *tienen el mismo carácter (atractor, repulsor, indiferente).*

(d) Si p *es un punto periódico de* f *con cuenca de atracción* $\mathcal{B}(p)$, *entonces la cuenca de atracción de* $\varphi(p)$ *es* $\varphi(\mathcal{B}(p))$.

(e) Los puntos periódicos de f *son densos en* D *si y sólo si los puntos periódicos de* g *son densos en* E.

(f) f *es caótica sobre* D *si y sólo si* g *es caótica sobre* E.

Demostración. El apartado *(a)* se sigue directamente de la definición y del hecho de que φ tiene inversa continua. Por consiguiente, se tiene que $f(\varphi^{-1}(u)) = \varphi^{-1}(g(u)), \quad u \in E$. Por otra parte, para $z \in D$ se tiene que

$$\varphi(f^2(z)) = \varphi(f(f(z))) = g(\varphi(f(z))) = g(g(\varphi(z))) = g^2(\varphi(z)).$$

Siguiendo un razonamiento inductivo, se deduce que

$$\varphi(f^n(z)) = \varphi(f(f^{n-1}(z))) = g(\varphi(f^{n-1}(z))) = g(g^{n-1}(\varphi(z))) = g^n(\varphi(z)),$$

con lo que se deduce el apartado *(b)*. Supongamos ahora que p es un punto periódico de f, de período k. Entonces k es el menor valor para el que $f^k(p) = p$. Teniendo en cuenta el apartado anterior, se tiene

$$\varphi(p) = \varphi(f^k(p)) = g^k(\varphi(p)). \tag{1.1}$$

En consecuencia $\varphi(p)$ es un punto periódico de g con el mismo período k, como se pone de manifiesto en *(c)*.

Para probar el apartado *(c')* tendremos en cuenta que, por el teorema 1.6 el carácter de un punto fijo p de f viene dado por el valor de $|f'(p)|$. Análogamente, el carácter de un punto periódico p, de período k viene dado por el valor de $|f'(p_0)f'(p_1)\cdots f'(p_{k-1})|$, donde $p_j = f^j(p)$ para $j = 0, 1, \ldots, k-1$. Teniendo esta notación en cuenta y derivando en la segunda igualdad de (1.1) tenemos que

$$\varphi'(f^k(p))f'(p_0)f'(p_1)\cdots f'(p_{k-1}) = g'(q_0)g'(q_1)\cdots g'(q_{k-1})\varphi'(p),$$

donde $q_j = g^j(\varphi(p))$ para $j = 0, 1, \ldots, k-1$. Como $\varphi'(f^k(p)) = \varphi'(p) \neq 0$, tenemos que el carácter de p como punto periódico de f y de $\varphi(p)$ como punto periódico de g es el mismo.

El apartado *(d)* se basa de nuevo en la idea de que la conjugación respeta las órbitas. Así, si la órbita de un punto $z_0 \in D$ converge a un punto periódico p de f, entonces, la órbita de $\varphi(z_0)$ converge al punto periódico $\varphi(p)$ de g. Recíprocamente, si la órbita de un punto $u_0 \in E$ converge a un punto periódico q de g, entonces, la órbita de $\varphi^{-1}(u_0)$ converge al punto periódico $\varphi^{-1}(q)$ de f. Por consiguiente, $z_0 \in \mathcal{B}(p)$ si y sólo si $\varphi(z_0) \in \varphi(\mathcal{B}(p))$.

Los apartados *(e)* y *(f)* hacen referencia a conceptos que introducimos en la sección 1.5. Para profundizar en estas ideas recomendamos al lector consultar dicha sección así como, entre otras, las referencias [43] y [91]. □

Mostramos a continuación unos cuantos ejemplos para ilustrar los resultados teóricos dados en el teorema anterior.

Ejemplo 1.10. *Las aplicaciones $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ definidas por $f(x) = 2x$ y $g(x) = 8x$ son conjugadas por el homeomorfismo $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ dado por $\varphi(x) = x^3$.*

Es una comprobación inmediata:

$$\varphi \circ f \circ \varphi^{-1}(x) = \varphi(2x^{1/3}) = 2^3(x^{1/3})^3 = 8x = g(x).$$

Ejemplo 1.11. *Las aplicaciones $f(x) = 2x$ y $g(x) = x/2$ no pueden ser conjugadas.*

En primer lugar, notemos que para todo $x > 0$,

$$\lim_{n\to\infty} f^n(x) = \lim_{n\to\infty} 2^n x = \infty.$$

Por otra parte, para todo $x \in \mathbb{R}$,

$$\lim_{n\to\infty} g^n(x) = \lim_{n\to\infty} \frac{x}{2^n} = 0.$$

Si f y g fuesen topológicamente conjugadas, existiría un homeomorfismo φ tal que $f^n = \varphi^{-1} \circ g^n \circ \varphi$. Entonces, para todo $x > 0$,

$$\lim_{n\to\infty} f^n(x) = \lim_{n\to\infty} \varphi^{-1} \circ g^n \circ \varphi(x) = \varphi^{-1}(0) \in \mathbb{R},$$

lo que es una contradicción. Por lo tanto, f y g no pueden ser conjugadas.

Ejemplo 1.12. *¿Qué relación debe existir entre los parámetros $a, b \in \mathbb{R}$ para que las aplicaciones cuadráticas $g_a(x) = ax(1-x)$ y $h_b(x) = 1-bx^2$ sean conjugadas por un homeomorfismo de la forma $\varphi(x) = \alpha x + \beta$, con $\alpha \neq 0$?*

Notemos en primer lugar que φ transforma el punto crítico de h_b en el punto crítico de g_a, es decir, $\varphi(0) = 1/2$. En consecuencia, $\beta = 1/2$ y, por tanto, $\varphi(x) = \alpha x + 1/2$. Para determinar α usamos la ecuación de conjugación

$$g_a \circ \varphi(x) = \varphi \circ h_b(x).$$

Tenemos entonces

$$\begin{aligned} g_a \circ \varphi(x) &= g_a\left(\alpha x + \frac{1}{2}\right) \\ &= a\left(\alpha x + \frac{1}{2}\right)\left(1 - \left(\alpha x + \frac{1}{2}\right)\right) \\ &= \frac{a}{4} - a\alpha^2 x^2. \end{aligned}$$

Por otra parte

$$\begin{aligned} \varphi \circ h_b(x) &= \varphi(1 - bx^2) \\ &= \alpha(1 - bx^2) + \frac{1}{2} \\ &= \alpha + \frac{1}{2} - \alpha b x^2. \end{aligned}$$

Igualando las partes constantes y los coeficientes que acompañan a x^2, obtenemos las ecuaciones

$$\begin{cases} a/4 = \alpha + 1/2 \\ a\alpha^2 = \alpha b. \end{cases}$$

De la segunda ecuación, obtenemos que $\alpha = b/a$. Reemplazando este valor en la primera de esas ecuaciones se obtiene que los parámetros a y b en esas dos familias de aplicaciones para los cuales tenemos la conjugación están relacionados por

$$b = \frac{a^2 - 2a}{4}.$$

En este caso, α debe ser elegido como

$$\alpha = \frac{b}{a} = \frac{a-2}{4}.$$

Por ejemplo, en el caso particular de que $a = 4$ y $b = 2$, se tiene que $\alpha = 1/2$. Por tanto, las aplicaciones $g_4(x) = 4x(1-x)$ y $h_2(x) = 1 - 2x^2$ son conjugadas por el homeomorfismo $\varphi(x) = (x+1)/2$.

El teorema 1.7 asegura que dos aplicaciones conjugadas topológicamente generan dinámicas equivalentes. Esta idea permite reducir el estudio de las dinámicas de algunas familias de funciones a algunos casos concretos. Uno de los ejemplos más significativos de esta situación resulta ser el estudio dinámico de los polinomios de segundo grado. Aunque, en principio un polinomio cuadrático puede depender de tres parámetros, tras un proceso de conjugación topológica se puede reducir el estudio a una familia de polinomios dependientes de un solo parámetro.

Ejemplo 1.13. *Sea $f(x) = ax^2 + bx + c$, con $a \neq 0$, un polinomio de segundo grado con coeficientes reales. Entonces, las dinámicas de esta familia de funciones son conjugadas topológicamente con las dinámicas de*

$$f_\mu(x) = x^2 + \mu, \quad \mu \in \mathbb{R}. \tag{1.2}$$

Es decir, de los tres parámetros que aparecen en f, dos de ellos son innecesarios para describir la dinámica, y en realidad sólo basta un parámetro, μ, para llevar a cabo esta tarea.

Afirmamos que de esos tres parámetros a, b y c dos de ellos son «ilusorios» (innecesarios) desde el punto de vista dinámico. Para verlo, consideremos una aplicación afín $L(x) = \alpha x + \beta$, con $\alpha \neq 0$. Tenemos que L es biyectiva y $L^{-1}(x) = (x - \beta)/\alpha$. La función $g(x) = (L^{-1} \circ f \circ L)(x)$ es también cuadrática y, por el teorema 1.7, la descripción de las órbitas de g equivale a la descripción de las órbitas de f. Recordemos que L es llamada una conjugación entre f y g, y como en este caso L es afín, la llamaremos conjugación afín. La expresión explícita de g es la siguiente

$$\begin{aligned} g(x) &= L^{-1} \circ f \circ L(x) \\ &= \frac{a\,(L(x))^2 + bL(x) + c - \beta}{\alpha} \\ &= a\alpha x^2 + (2a\beta + b)x + \frac{a\beta^2 + b\beta + c - \beta}{\alpha}. \end{aligned}$$

Ahora, elegimos α y β de modo que $\alpha a = 1$ y $2a\beta + b = 0$, es decir, $\alpha = 1/a$ y $\beta = -b/(2a)$. De aquí resulta que $g(x) = x^2 + \mu$, con

$$\mu = \frac{4ac + 2b - b^2}{4}.$$

Finalmente, denotamos g por f_μ, pues depende sólo del parámetro μ.

Notemos que los polinomios definidos en (1.2) tienen puntos fijos en los valores

$$\frac{1 \pm \sqrt{1 - 4\mu}}{2},$$

que son reales para $\mu \leq 1/4$ y complejos para $\mu > 1/4$. Si nos centramos en los números reales, esto hace que el comportamiento dinámico de los polinomios de la familia definida en (1.2) sea sencillo para $\mu > 1/4$ pero mucho más complicado para $\mu \leq 1/4$. Vamos a precisar estos aspectos en los dos siguientes ejemplos.

Ejemplo 1.14. *Si $\mu > 1/4$, los iterados de cualquier punto por f_μ tienden a $+\infty$. Si $\mu = 1/4$, se tiene que $p = 1/2$ es un punto fijo indiferente con cuenca de atracción $[-1/2, 1/2]$.*

El primer caso es de comprobación inmediata, pues f_μ no tiene puntos fijos y su gráfica está por encima de la diagonal $y = x$. En el segundo caso, si $\mu = 1/4$, aparece un único punto fijo en $x = 1/2$. Además, $f_{1/4}(-1/2) = 1/2$, y los puntos en el intervalo $[-1/2, 1/2]$ se aproximan a $1/2$ por iteraciones de $f_{1/4}$. Por otra parte, los puntos de $(-\infty, -1/2) \cup (1/2, \infty)$ tienden a $+\infty$ bajo iteraciones de $f_{1/4}$.

Ejemplo 1.15. *Para $\mu \leq 1/4$, las dinámicas de la familia cuadrática $f_\mu(x) = x^2 + \mu$ son conjugadas topológicamente con las dinámicas de la función logística*

$$\ell_\lambda(x) = \lambda x(1 - x).$$

Consideremos una transformación afín $T(x) = \gamma x + \delta$, con $\gamma \neq 0$ y calculemos $T^{-1} \circ f_\mu \circ T(x)$:

$$\begin{aligned} T^{-1} \circ f_\mu \circ T(x) &= \frac{f_\mu\left(T(x)\right) - \delta}{\gamma} \\ &= \frac{\left(T(x)\right)^2 + \mu - \delta}{\gamma} \\ &= \gamma x^2 + 2\delta x + \frac{\delta^2 + \mu - \delta}{\gamma}. \end{aligned}$$

Ahora hacemos $\delta^2 - \delta + \mu = 0$, es decir,

$$\delta = \frac{1 \pm \sqrt{1 - 4\mu}}{2}.$$

Notemos que es aquí donde se requiere la condición $\mu \leq 1/4$. Usamos sólo el valor positivo de la raíz, es decir, $\delta = (1 + \sqrt{1 - 4\mu})/2$ y reemplazando obtenemos:

$$T^{-1} \circ f_\mu \circ T(x) = \gamma x^2 + (1 + \sqrt{1 - 4\mu})x.$$

Tomando ahora $\gamma = -(1 + \sqrt{1 - 4\mu})$, resulta que

$$T^{-1} \circ f_\mu \circ T(x) = \gamma x(1 - x)$$

que es la función logística con $\lambda = \gamma$.

1.4. Estudio dinámico de la función logística

En la sección anterior hemos visto que las dinámicas de la función logística

$$\ell_\lambda(x) = \lambda x(1-x). \tag{1.3}$$

pueden servir para estudiar el comportamiento dinámico de las funciones definidas en (1.2) y, por lo tanto, de cualquier polinomio de segundo grado. En concreto, la relación entre los parámetros μ y λ que aparecen en las funciones definidas en (1.2) y (1.3) es

$$\lambda = 1 + \sqrt{1-4\mu}.$$

Notemos que la situación que nos queda pendiente de estudio para la familia (1.2) es $\mu < 1/4$, que conduce a estudiar la función logística (1.3) para $\lambda > 1$. No obstante, vamos a proceder a realizar un estudio más general de las dinámicas de (1.3) para $\lambda \geq 0$.

La función logística (1.3) es un clásico entre los estudios de las dinámicas de funciones de variable real. A pesar de su aparente sencillez, en ella se engloban la mayoría de situaciones que se nos pueden plantear en el estudio de una función cualquiera, incluyendo fenómenos tales como las bifurcaciones o comportamientos caóticos. Se pueden encontrar un gran número de referencias sobre la función logística. Como una pequeña muestra, se pueden consultar [15], [43], [56], [93], [97] o [114].

Antes de comenzar con el estudio dinámico de la función logística, observemos que podemos reducir el intervalo de estudio al intervalo $I = [0,1]$. En efecto, si $x_0 \notin I$, entonces $x_1 = \ell_\lambda(x_0) < 0$ y el resto de iteraciones por ℓ_λ van decreciendo hacia $-\infty$.

Notemos también que para cada λ se tiene que $\ell_\lambda(0) = 0$ y $\ell_\lambda(1) = 0$. Además, la parábola definida por ℓ_λ tiene un máximo absoluto cuando $x = 1/2$.

Por último, observemos que ℓ_λ tiene dos puntos fijos: 0 y $x_\lambda = (\lambda-1)/\lambda$. De la ubicación del segundo punto fijo y del carácter atractor o repulsor de ambos va a depender en gran manera el comportamiento dinámico de la función logística (1.3). Nótese que

$$\ell'_\lambda(0) = \lambda \quad \text{y} \quad \ell'_\lambda\left(\frac{\lambda-1}{\lambda}\right) = 2-\lambda. \tag{1.4}$$

De (1.4) se obtienen las caracterizaciones de los puntos fijos de (1.3) que se muestran en el cuadro 1.1.

Después de estas consideraciones generales, vamos a pasar al estudio de la función logística (1.3) para distintos valores de λ y para $x \in I = [0,1]$.

La función logística para $0 \leq \lambda < 1$

Dejando a un lado el caso $\lambda = 0$ para el cual la función logística tiene un comportamiento trivial (de hecho, la órbita de cualquier punto $x \in I$ es $\{x, 0, 0 \dots\}$), se tiene que ℓ_λ sólo tiene

Cuadro 1.1: Comportamiento de los puntos fijos de la función logística, 0 y $x_\lambda = (\lambda - 1)/\lambda$, para diferentes valores de λ. Notemos que para $\lambda = 1$ ambos puntos fijos coinciden.

	$0 \leq \lambda < 1$	$\lambda = 1$	$1 < \lambda < 3$	$\lambda = 3$	$\lambda > 3$
0	Atractor	Indiferente	Repulsor	Repulsor	Repulsor
x_λ	Fuera de I	Indiferente	Atractor	Indiferente	Repulsor

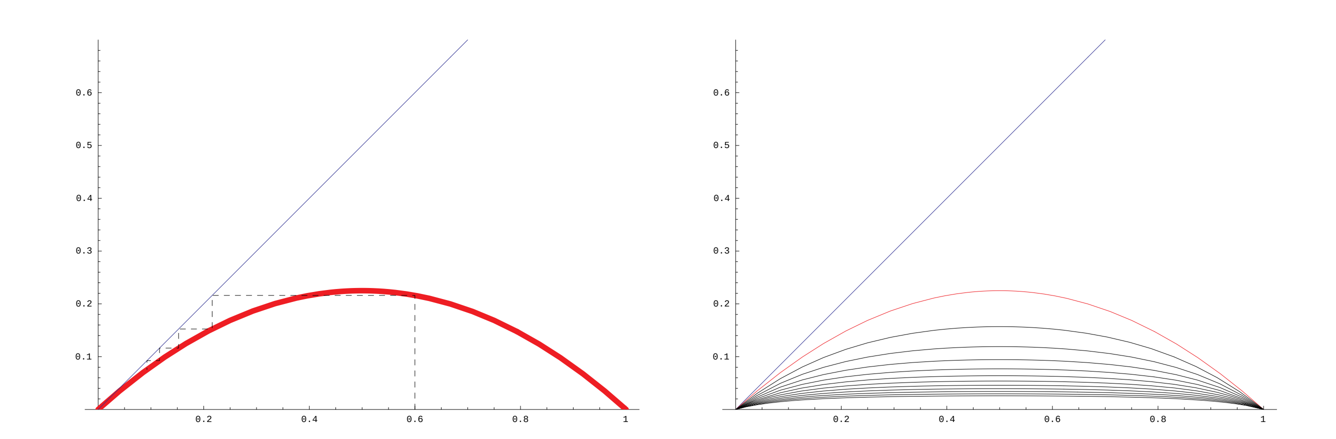


Figura 1.4: Función logística para $\lambda = 0.9$: órbita de $x_0 = 0.6$ y comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0, 1]$.

en el intervalo I a 0 como punto fijo. Además este punto fijo es atractor ya $\ell'_\lambda(0) = \lambda < 1$ (véase (1.4)). Además la órbita de cualquier punto $x_0 \in (0, 1)$ decrece a 0 cuando $n \longrightarrow \infty$.

En la figura 1.4 se muestra, para $\lambda = 0.9$, la órbita del punto $x_0 = 0.6$ y el comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0, 1]$ que se obtiene representando las gráficas de las composiciones iteradas de la función ℓ_λ consigo misma. Como se aprecia en la figura estas iteraciones se aproximan a la función constante igual a cero.

La función logística para $\lambda = 1$

Para $\lambda = 1$, los dos posibles puntos fijos de la función logística, 0 y $x_\lambda = (\lambda - 1)/\lambda$, confluyen en uno sólo: $x = 0$. En este caso, se trata de un atractor «débil» en $[0, 1]$, ya que, $\ell'_1(0) = 1$. En este caso, la órbita de cualquier punto $x_0 \in (0, 1)$ también decrece a 0 cuando $n \longrightarrow \infty$. El comportamiento gráfico es muy similar al que se muestra en la figura 1.4.

La función logística para $1 < \lambda \leq 2$

A partir de ahora, para $\lambda > 1$, se tiene que $x = 0$ es punto fijo repulsor. Pero, por otra parte, el otro punto fijo $x_\lambda = (\lambda - 1)/\lambda$ está en el interior del intervalo I. Además, en el caso $1 < \lambda < 3$, se trata de un punto fijo atractor.

En efecto, para $1 < \lambda \leq 2$, se sigue de (1.4) que $0 \leq \ell'_\lambda(x_\lambda) < 1$. Además, en este caso, se puede probar que la función logística tiene únicamente dos puntos eventualmente fijos: $x = 1$ y $x = 1/\lambda$ (nótese que para $\lambda = 2$ este segundo punto coincide con el punto fijo atractor x_λ, que ahora resulta ser un superatractor). Los puntos eventualmente fijos tienen su influencia en el comportamiento dinámico de la función logística. En concreto se tiene:

- Las órbitas de $x_0 \in (0, x_\lambda)$ crecen monótonamente hacia x_λ.
- Las órbitas de $x_0 \in (x_\lambda, 1/\lambda)$ decrecen monótonamente hacia x_λ.
- Si $x_0 \in (1/\lambda, 1)$ se tiene que $x_1 = \ell_\lambda(x_0) \in (0, x_\lambda)$. A partir de aquí, las órbitas crecen monótonamente hacia x_λ.

Estos comportamientos se muestran, para algunos casos concretos, en la figura 1.5.

En este caso, las funciones iteradas ℓ_λ^n tienden a las aplicaciones constantes dadas por $y = x_\lambda$, como se muestra en la figura 1.6.

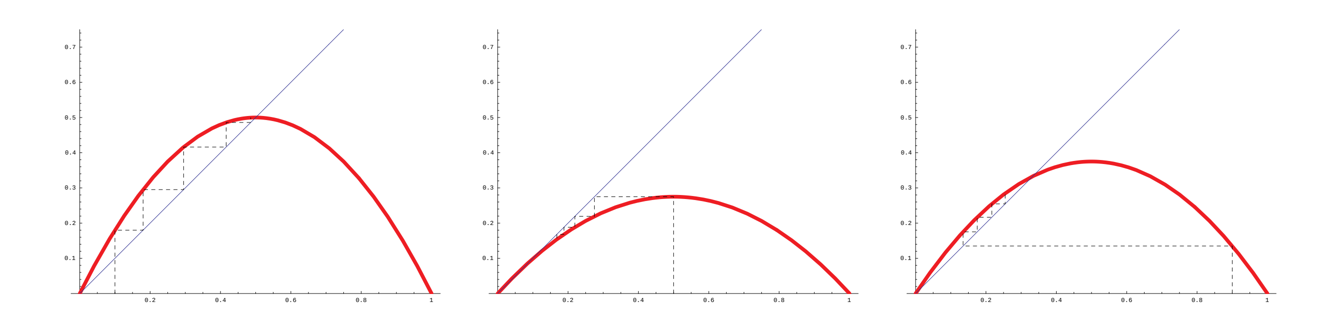


Figura 1.5: Órbitas de $x_0 = 0.1$ para $\lambda = 2$, de $x_0 = 0.5$ para $\lambda = 1.1$ y de de $x_0 = 0.9$ para $\lambda = 1.5$. Nótese la distinta forma de converger a los respectivos puntos fijos 1/2, 1/11 y 1/3.

La función logística para $2 < \lambda < 3$

El carácter de los puntos fijos es el mismo que en el caso anterior: $x = 0$ es repulsor y $x_\lambda = (\lambda - 1)/\lambda$ es atractor, ya que de (1.4) se deduce $-1 < \ell'_\lambda(x_\lambda) < 0$. La principal novedad es que ahora aparecen una cantidad numerable de puntos eventualmente fijos. Estos puntos tienen su influencia en el comportamiento dinámico de la función logística. En concreto, en este caso, las órbitas se aproximan por ambos lados a x_λ y aparece una «espiral» en el seguimiento de la órbita, como puede verse en la figura 1.7.

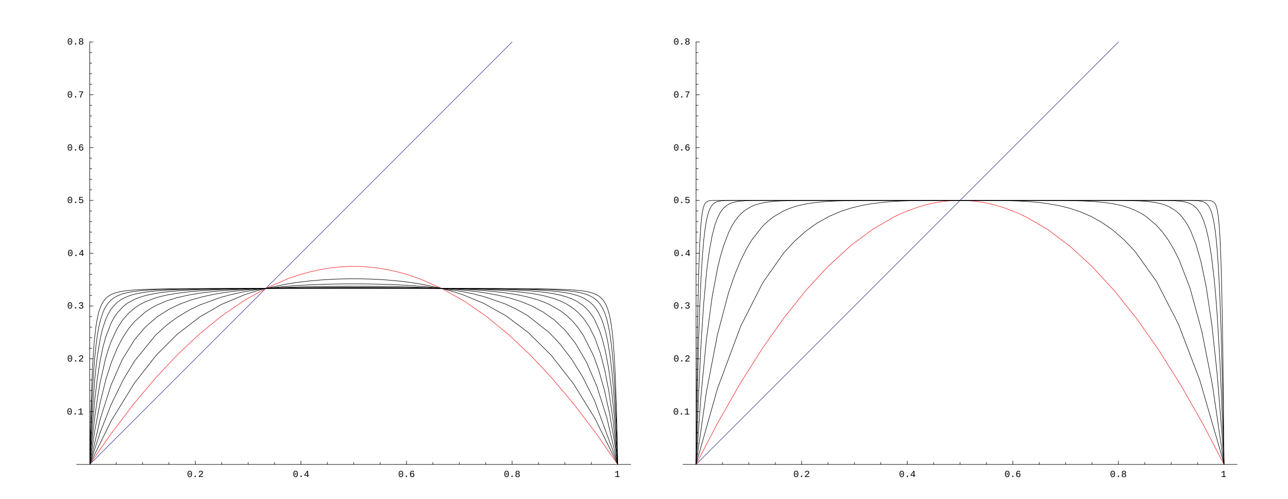


Figura 1.6: Comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0,1]$ para $\lambda = 1.5$ y para $\lambda = 2$.

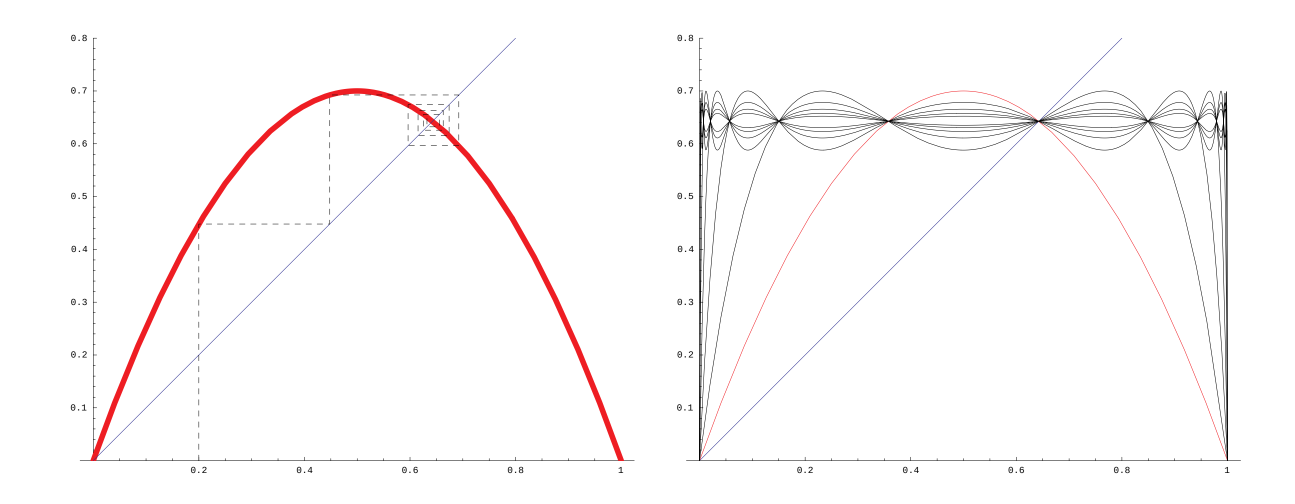


Figura 1.7: Función logística para $\lambda = 2.8$: órbita de $x_0 = 0.2$ y comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0,1]$.

La función logística para $\lambda = 3$

El comportamiento en este caso es muy similar al caso anterior, con la única diferencia de que el punto fijo $x_\lambda = 2/3$ es un punto indiferente en lugar de atractor. No obstante su cuenca de atracción sigue siendo todo el intervalo $(0,1)$. El comportamiento gráfico es similar

al mostrado en la figura 1.7, aunque con una convergencia mucho más lenta.

La función logística para $3 < \lambda \leq 1+\sqrt{6} \approx 3.4495$

Para $\lambda > 3$, ambos puntos fijos 0 y x_λ son repulsores. ¿Qué ocurre con las órbitas de los puntos de I, distintos de 0 y x_λ?

Considerando la aplicación $\ell_\lambda^2 = \ell_\lambda \circ \ell_\lambda$, vemos que su gráfico intersecta la diagonal en otros dos puntos aparte de los puntos fijos de ℓ_λ. Esto significa que existen puntos x_1, x_2 tales que $\ell_\lambda(x_1) = x_2$ y $\ell_\lambda(x_2) = x_1$, en otras palabras, aparece una órbita periódica de período 2, también denominada un 2-ciclo. Podemos observar esta situación en las gráficas de la figura 1.8.

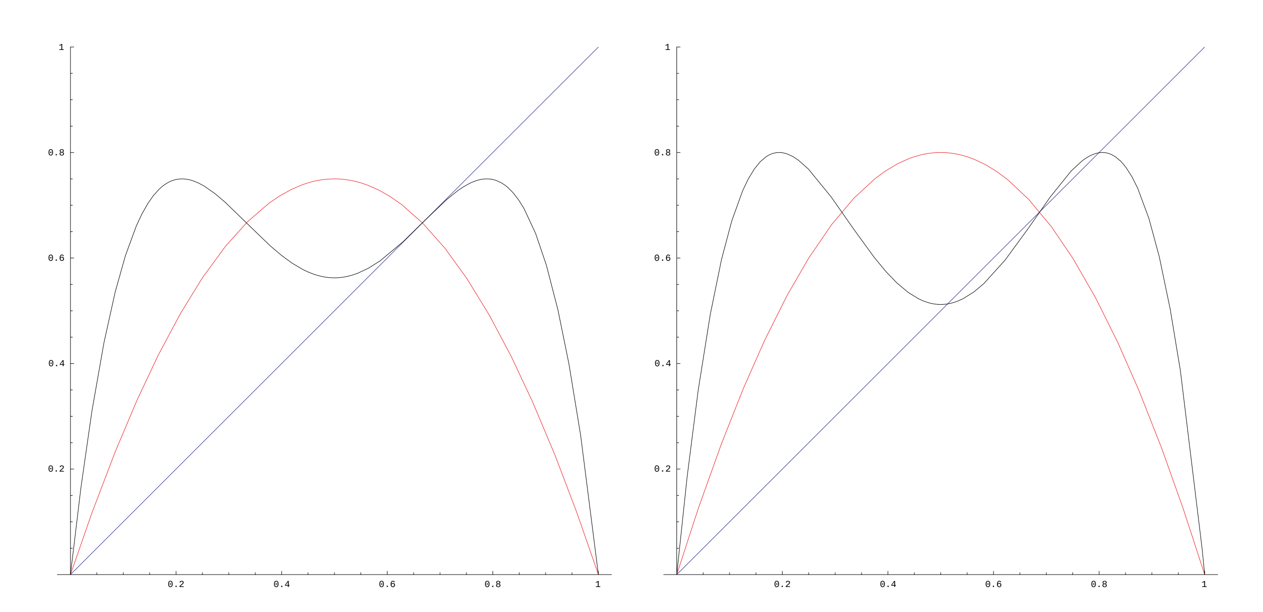


Figura 1.8: En negro, las gráficas de ℓ_λ^2 para $\lambda = 3$ y para $\lambda = 3.2$: el punto fijo x_λ se ha bifurcado en dos puntos fijos de ℓ_λ^2.

Lo que ha ocurrido es que el punto fijo x_λ de ℓ_λ que se comportaba como un atractor para $1 < \lambda \leq 3$, se ha descompuesto en el ciclo atractor $\{x_1, x_2\}$. Para encontrar este ciclo, analizamos las soluciones de la ecuación

$$x = \ell_\lambda\left(\ell_\lambda(x)\right) = \lambda\left(\lambda x(1-x)\right)\left(1 - \lambda x(1-x)\right).$$

Ésta es una ecuación de cuarto grado cuyas raíces son los puntos fijos, $x = 0$ y x_λ, además de los puntos sobre la órbita periódica. Encontramos x_1 y x_2, dividiendo el polinomio por $x\left(x - x_\lambda\right)$. Así, después de unas pocas manipulaciones algebraicas, se llega a la expresión

$$\frac{\ell_\lambda\left(\ell_\lambda(x)\right) - x}{x\left(x - x_\lambda\right)} = -\lambda(\lambda^2 x^2 - \lambda^2 x - \lambda x + \lambda + 1).$$

Las soluciones de esta ecuación cuadrática son:

$$x_1, x_2 = \frac{1}{2} + \frac{1}{2\lambda} \pm \frac{1}{2}\sqrt{\left(1+\frac{1}{\lambda}\right)\left(1-\frac{3}{\lambda}\right)}.$$

La estabilidad del ciclo depende del valor $(\ell_\lambda^2)'(x_j)$, $j = 1, 2$, el cual es el mismo para x_1 y x_2. Notemos que

$$\left(\ell_\lambda^2\right)'(x_j) = \ell_\lambda'\left(\ell_\lambda(x_j)\right) \cdot \ell_\lambda'(x_j) = \ell_\lambda'(x_1) \cdot \ell_\lambda'(x_2).$$

En consecuencia, $|(\ell_\lambda^2)'(x_j)| < 1$ si $|4 + 2\lambda - \lambda^2| < 1$. Resolviendo esta desigualdad para $0 \leq \lambda \leq 4$, encontramos que es cierta para $3 < \lambda \leq 1 + \sqrt{6} \approx 3.449489743$. Entonces, para λ en este intervalo, la órbita de cualquier punto del intervalo, excepto los puntos fijos y los eventualmente fijos, tiende al 2-ciclo atractor $\{x_1, x_2\}$.

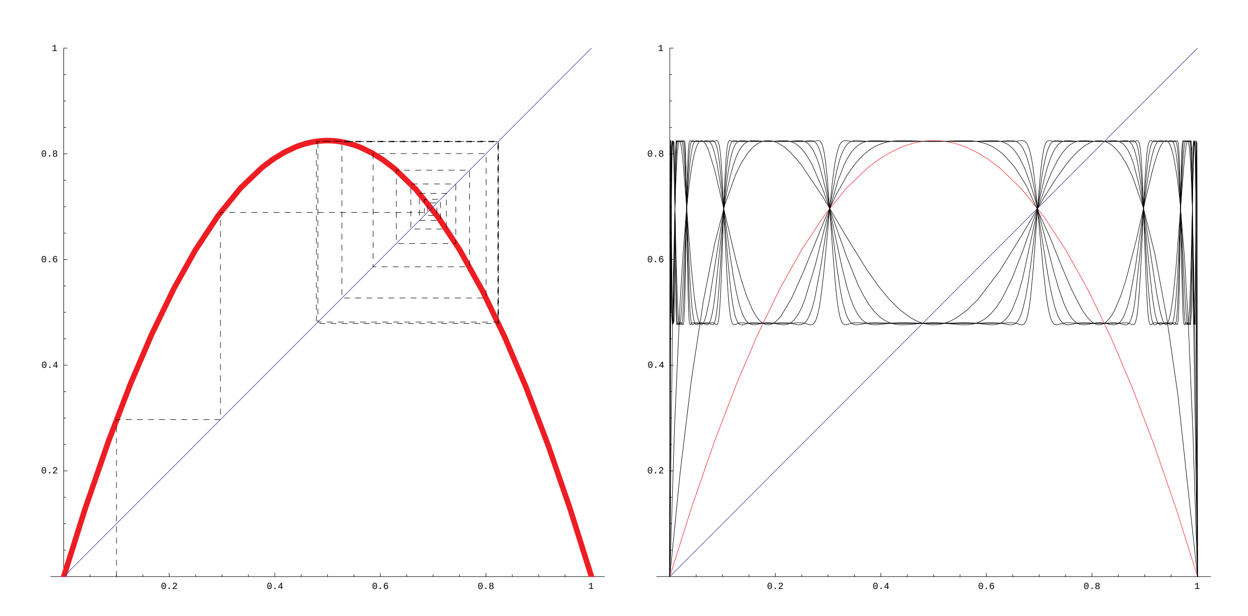


Figura 1.9: Función logística para $\lambda = 3.3$: órbita de $x_0 = 0.1$ y comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0, 1]$. En ambos casos, se puede observar la presencia de un 2-ciclo atractor.

La función logística para $1 + \sqrt{6} < \lambda \leq \lambda_\infty \approx 3.5699$

Ahora, para $\lambda > 1 + \sqrt{6} \approx 3.4495$ la órbita periódica $\{x_1, x_2\}$ es repulsora y en su lugar aparece una órbita atractora de período 4. Esta órbita puede encontrarse resolviendo numéricamente $\ell_\lambda^4(x) = x$ y $\ell_\lambda^2(x) \neq x$. En concreto, se puede mostrar que esa órbita periódica de período 4 es atractora para $3.4495\ldots < \lambda < 3.5441\ldots$ y repulsora para $\lambda > 3.5441\ldots$

Para $\lambda > 3.5441\ldots$ una órbita periódica atractora de período 8 emerge. Numéricamente esta se obtiene resolviendo $\ell_\lambda^8(x) = x$ y $\ell_\lambda^4(x) \neq x$. Esta órbita de período 8 es atractora para $3.5441\ldots < \lambda < 3.54644\ldots$ y repulsora para $\lambda > 3.54644\ldots$

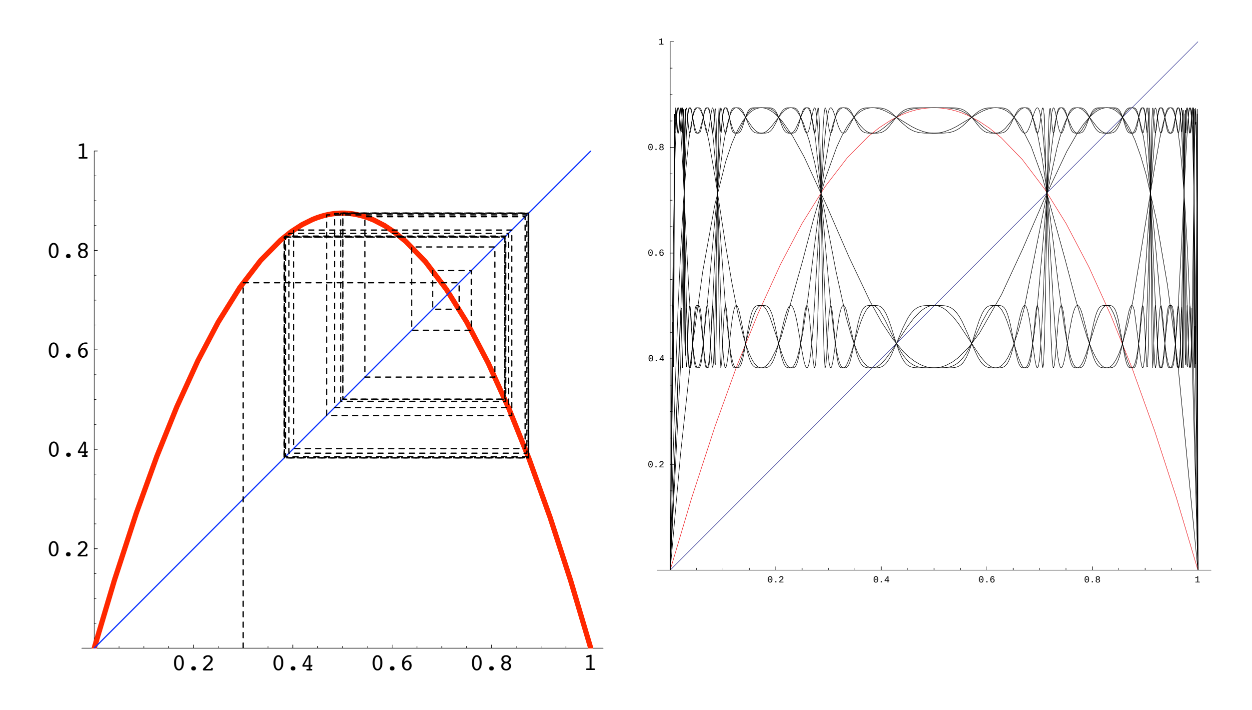


Figura 1.10: Función logística para $\lambda = 3.5$: órbita de $x_0 = 0.3$ y comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0, 1]$. En ambos casos, puede observarse la presencia de un 4-ciclo atractor.

Observemos ahora los valores para los cuales apareció una nueva órbita periódica, de período el doble de la que había aparecido antes: $\lambda_0 = 1$, $\lambda_1 = 3$, $\lambda_2 = 1 + \sqrt{6} \approx 3.4495\ldots$, $\lambda_3 \approx 3.5441\ldots$, $\lambda_4 \approx 3.5644\ldots$ Llamamos a este fenómeno *bifurcación de duplicación de período*. Puede probarse que este fenómeno se va repitiendo, pasando por todos los períodos de la forma 2^n, hasta llegar a un valor límite

$$\lambda_\infty = \lim_{n\to\infty} \lambda_n \approx 3.5699.$$

Al valor λ_∞ se le llama *constante de Feigenbaum* o *punto de entrada al caos*. En este caso existe una familia numerable de órbitas periódicas repulsoras correspondientes a todos los períodos de la forma 2^n. Pero además, existe un conjunto invariante de tipo Cantor y con carácter atractor contenido en el complementario de los ciclos repulsores.

El diagrama de bifurcación para la función logística, conocido como *diagrama de Feigenbaum* se muestra en la figura 1.11. Básicamente se construye representando los puntos a los que converge la órbita del punto $x_0 = 1/2$ para diferentes valores del parámetro λ. En el diagrama se aprecian las diferentes duplicaciones del período que se van produciendo y cómo se vuelve extremadamente intrincado a partir del punto λ_∞. En la figura 1.12 se muestran dos partes ampliadas del diagrama de Feigenbaum en las que se aprecia las distintas bifurcaciones que se van produciendo.

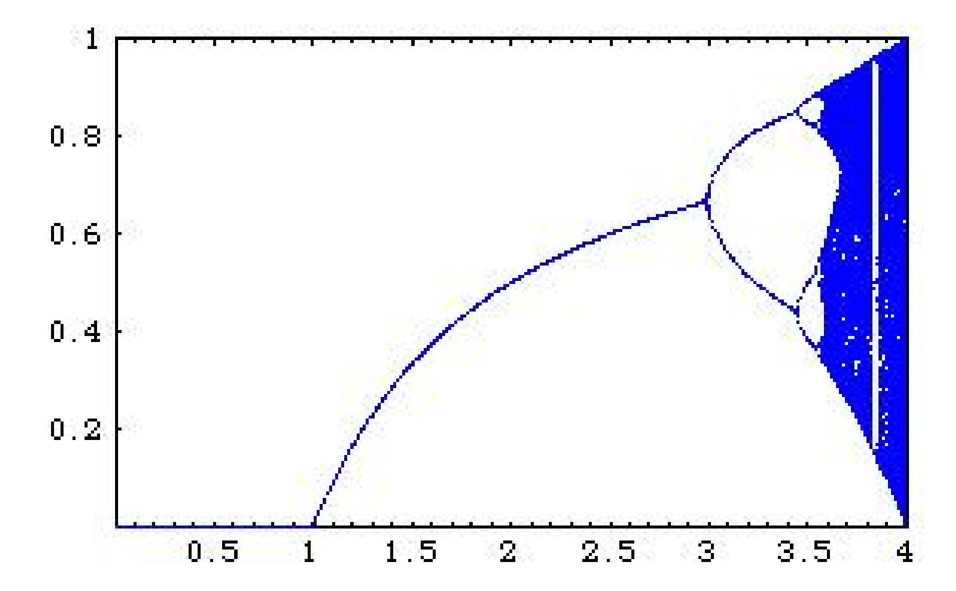


Figura 1.11: Diagrama de Feigenbaum asociado a la función logística $\ell_\lambda(x)$ definida en (1.3). En el eje de abcisas se muestran los valores del parámetro λ y en el eje de ordenadas los puntos a los que converge la órbita del punto $x_0 = 1/2$.

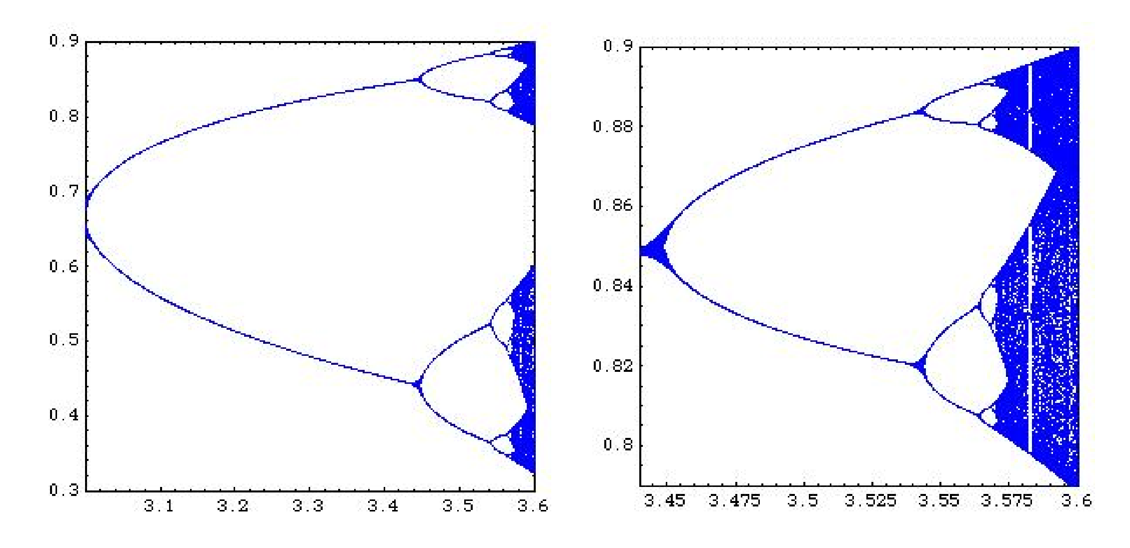


Figura 1.12: Dos ampliaciones del diagrama de Feigenbaum.

La función logística para $\lambda_\infty < \lambda < 4$

El punto λ_∞ marca la separación entre el régimen de períodos y el régimen caótico para esta familia cuadrática: las órbitas de dos puntos muy cercanos pueden ser totalmente distintas. Esto se observa por ejemplo analizando las series temporales de $x_0 = 0.75$ y $x_0 = 0.75000001$ para $\lambda = 4$.

Aunque para $\lambda > \lambda_\infty$, la mayoría de los casos presentan un comportamiento caótico, todavía existen algunos valores de λ con un comportamiento periódico, dando lugar a unas franjas de regularidad dentro del diagrama de Feigenbaum conocidas como *islas de estabilidad.*

- Para $\lambda \approx 3.627$ aparece un 6-ciclo atractor.

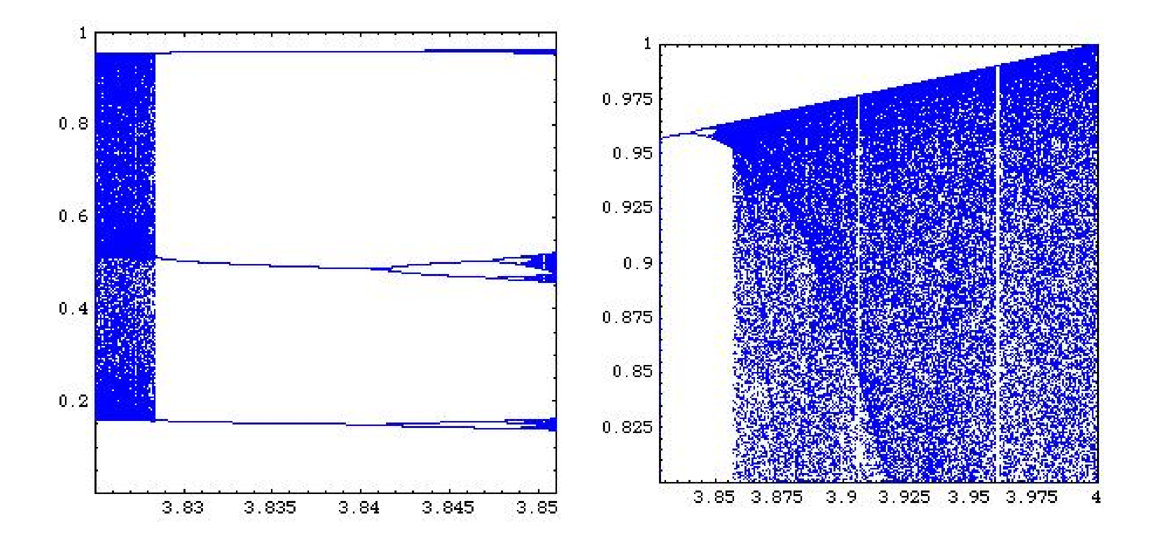


Figura 1.13: Dos nuevas ampliaciones del diagrama de Feigenbaum: en la primera de ellas se aprecia la aparición de un 3-ciclo atractor que se va bifurcando, dando lugar a una isla de regularidad. El segundo detalle muestra el intrincado comportamiento para valores del parámetro λ cercanos a 4.

- Para $\lambda \approx 3.702$ aparece un 7-ciclo atractor.
- Para $\lambda \approx 3.74$ aparece un 5-ciclo atractor.
- Para $\lambda = 1 + \sqrt{8} \approx 3.828$ aparece un 3-ciclo atractor, que a su vez se irá bifurcando en ciclos de períodos $6, 12, 24, \ldots$ hasta llegar a un valor límite aproximadamente 3.855 (véase la figura 1.13).

La función logística para $\lambda = 4$

En este caso, existen puntos periódicos repulsores de todos los períodos posibles. Además, éstos son densos en el intervalo $I = [0, 1]$, es decir, dado un intervalo abierto cualquiera en I, siempre existe un punto periódico en su interior. Como se puede ver en la figura 1.14, el comportamiento de las órbitas de un punto $x_0 \in I$ es caótico y la evolución de las gráficas de las composiciones de la función ℓ_4 consigo misma tiende a «rellenar» el cuadrado $[0, 1] \times [0, 1]$.

La función logística para $\lambda > 4$

Al igual que en el caso $\lambda = 4$, existen puntos periódicos repulsores de todos los períodos posibles. Como novedad, a partir de $\lambda = 4$, aparecen puntos cuyas órbitas escapan del intervalo $I = [0, 1]$ y tienen un comportamiento divergente a $-\infty$. La órbita de uno de estos puntos se muestra en la figura 1.15, junto con la evolución de las gráficas de las composiciones de la función ℓ_λ consigo misma. Como se aprecia, las gráficas de dichas composiciones escapan ya del intervalo $I = [0, 1]$.

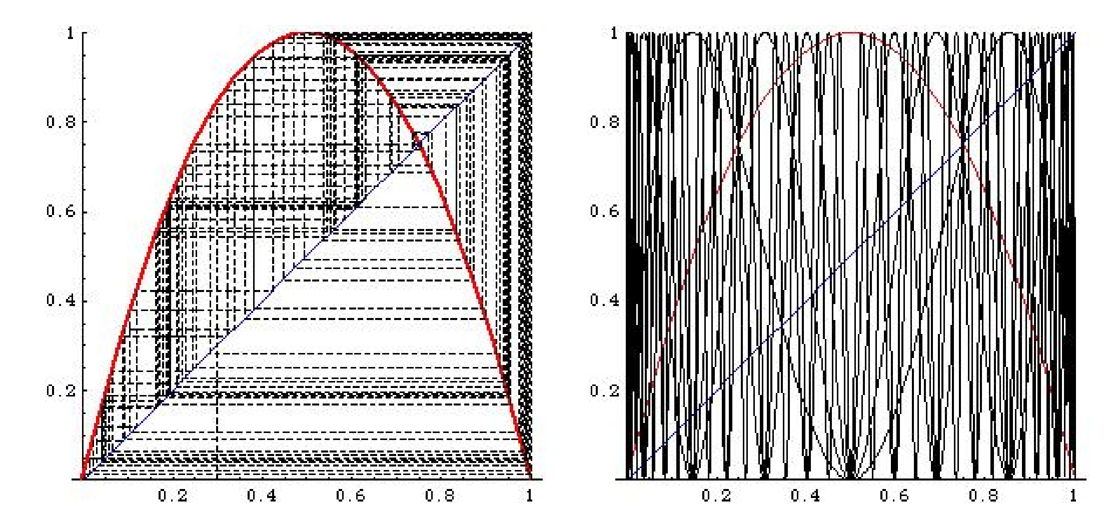


Figura 1.14: Función logística para $\lambda = 4$: órbita de $x_0 = 0.3$ y comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0, 1]$.

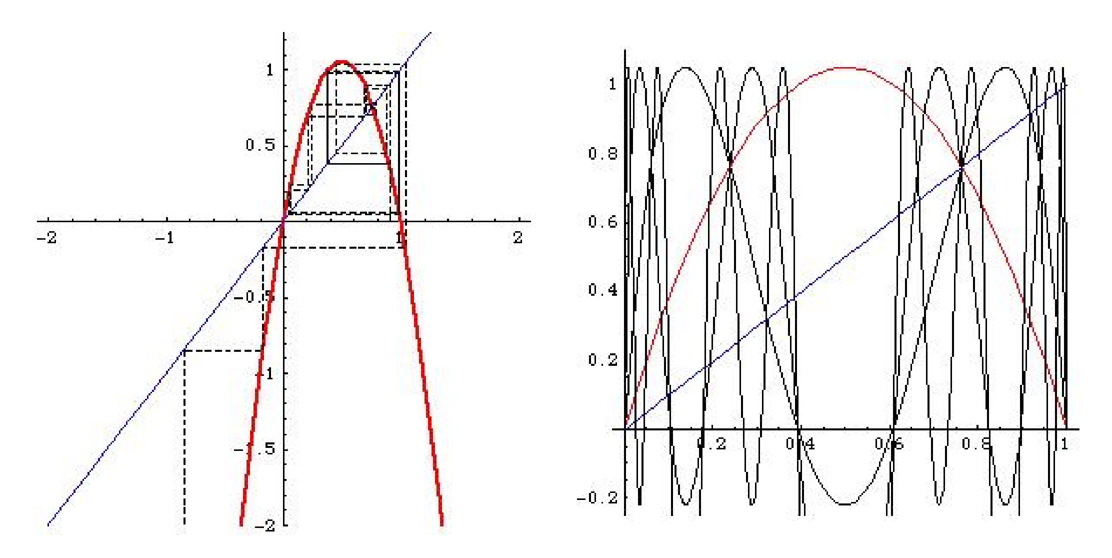


Figura 1.15: Función logística para $\lambda = 4.2$: órbita de $x_0 = 0.9$ y comportamiento conjunto de las órbitas de todos los puntos del intervalo $[0, 1]$.

Series temporales de la función logística

Para finalizar esta sección, y a modo de resumen, analizamos las *series temporales* de la función logística para distintos valores del parámetro λ. En las figuras 1.16–1.19 se muestran las gráficas de las órbitas de un mismo punto, $x_0 = 0.75$, para algunos valores representativos del parámetro λ. En estas figuras el número de iteraciones n aparece en el eje horizontal, mientras que en el eje vertical se representan los correspondientes valores de $x_n = \ell^n(x_0)$.

En las cuatro primera figuras, correspondientes a los valores de los parámetros $\lambda = 0.8$, $\lambda = 2$, $\lambda = 2.8$ y $\lambda = 3$, se aprecia la convergencia hacia un punto fijo. Las tres figuras siguientes muestran la aparición de ciclos periódicos. En concreto un 2-ciclo para $\lambda = 3.3$, un 4-ciclo para $\lambda = 3.5$ y un 8-ciclo para $\lambda = 3.6$. Las dos últimas series temporales de la figura 1.18 muestran ya una dinámica mucho más complicada.

Finalmente, en la figura 1.19, se aprecia ya el comportamiento caótico y una gran sensibilidad a las condiciones iniciales. De hecho, pequeños cambios en los datos iniciales provocan unos resultados totalmente diferentes.

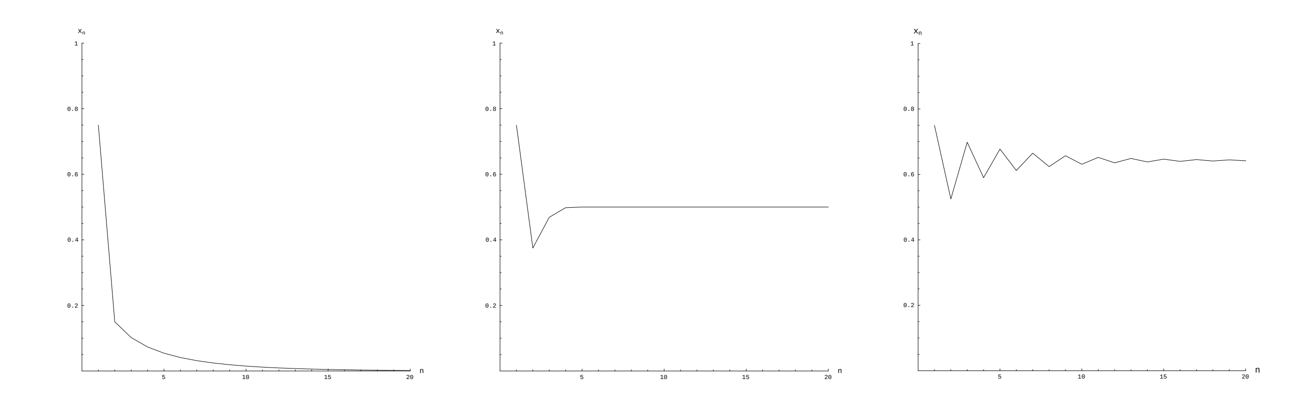


Figura 1.16: Series temporales de la función logística para $x_0 = 0.75$, con $\lambda = 0.8$, $\lambda = 2$ y $\lambda = 2.8$.

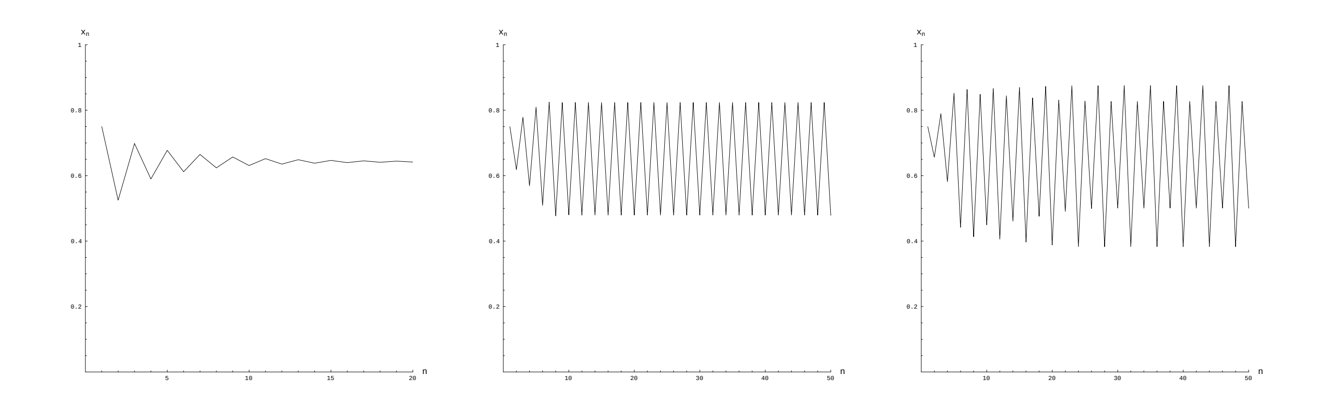


Figura 1.17: Series temporales de la función logística para $x_0 = 0.75$, con $\lambda = 3$, $\lambda = 3.3$ y $\lambda = 3.5$.

1.5. Sistemas dinámicos caóticos

Con el estudio dinámico de la función logística ha aparecido de forma experimental el concepto de *caos*. Llegados a este punto, se hace necesario definir dicho concepto con más precisión. Para ello, enunciamos a continuación la definición de caos introducida por Devaney en 1992 [42]. Previamente, necesitamos definir tres nuevos conceptos.

Definición 1.11. *Un sistema dinámico discreto* (X, f) *se dice* topológicamente transitivo *si dados dos subconjuntos abiertos cualesquiera* U *y* V *de* X*, existe* $n \in \mathbb{N}$ *tal que* $f^n(U) \cap V \neq \emptyset$.

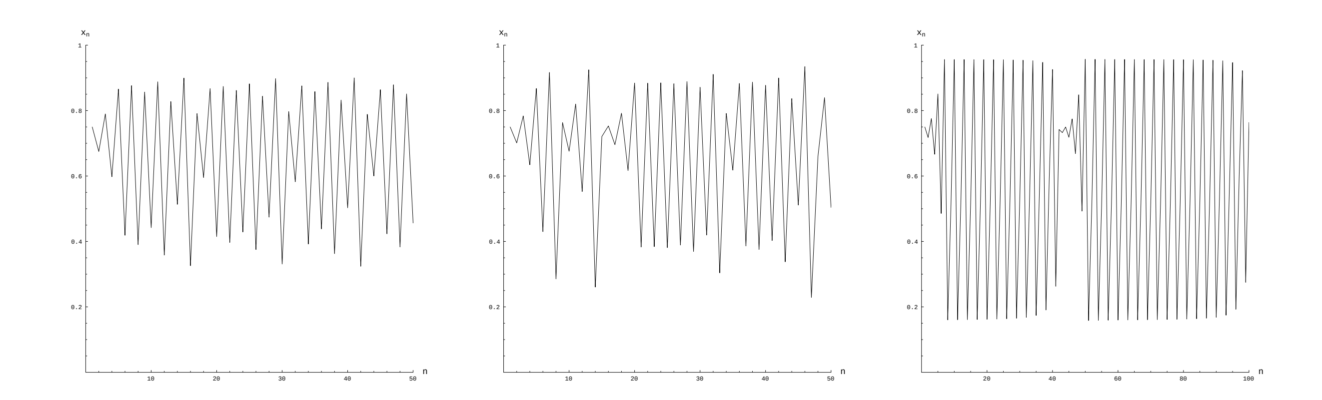


Figura 1.18: Series temporales de la función logística para $x_0 = 0.75$, con $\lambda = 3.6$, $\lambda = 3.74$ y $\lambda = 3.828$.

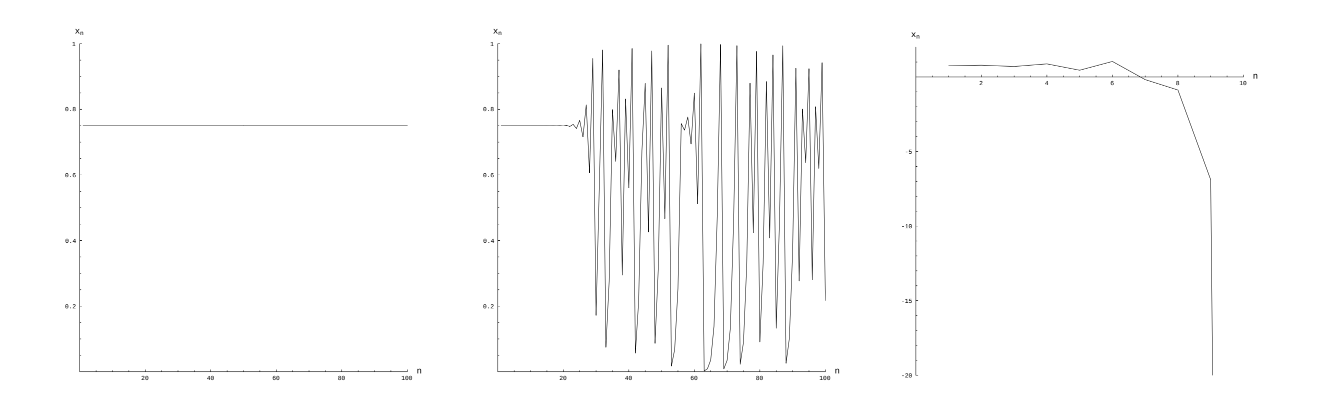


Figura 1.19: Series temporales de la función logística para $x_0 = 0.75$, con $\lambda = 4$, $x_0 = 0.750000001$, con $\lambda = 4$ y $x_0 = 0.75$, con $\lambda = 4.2$.

Definición 1.12. *Sea Y un subconjunto de X. Se dice que Y es* denso *en X si para cualquier subconjunto abierto U de X, siempre existe un punto de Y en U.*

Definición 1.13. *Un sistema dinámico discreto (X, f) es* sensible respecto a las condiciones iniciales *si existe un $\delta > 0$ tal que, para todo $x \in X$ y para todo $\varepsilon > 0$, existen $y \in X$ y $n \in \mathbb{N}$ tales que*

$$d(x, y) < \varepsilon \quad y \quad d(f^n(x), f^n(y)) > \delta,$$

donde d denota la distancia correspondiente en el espacio métrico X. A este fenómeno se le conoce también con el nombre de efecto mariposa.

Definición 1.14. *Un sistema dinámico discreto (X, f) se dice* caótico *si cumple las tres condiciones siguientes:*

1. *Es topológicamente transitivo.*

2. *El conjunto de puntos periódicos de f es denso en X.*

3. *El sistema es sensible respecto a las condiciones iniciales.*

Con posterioridad a la definición de Devaney, en 1992, Banks et al. [10] demuestran que la sensibilidad a las condiciones iniciales es una exigencia redundante, ya que se puede deducir de las otras dos propiedades. No obstante, eliminar la sensibilidad de las condiciones iniciales de la definición, nos proporciona una idea menos intuitiva del concepto de caos. Aún podemos encontrar una caracterización más del concepto de sistema dinámico caótico, dada por Touhey [145] en 1997 que establece que un sistema dinámico es caótico si y sólo si para cualesquiera U y V conjuntos abiertos de X, existe una órbita periódica que visita ambos. Esta propiedad implica las tres condiciones de la definición de Devaney.

Teorema 1.8 (Touhey, [145])**.** *Sea (X, f) un sistema dinámico discreto tal que para dos conjuntos abiertos no vacíos cualesquiera $U, V \subseteq X$, existe un punto periódico $p \in U$ tal que $f^n(p) \in V$ para algún $n \in \mathbb{N}$, entonces el sistema dinámico es caótico en el sentido de la definición de Devaney.*

Veamos ahora unos ejemplos básicos de sistemas dinámicos caóticos.

Ejemplo 1.16. *El sistema dinámico (X, S), donde $X = [0, 1]$ y S es la función conocida como «diente de sierra» (*saw-tooth function*) y definida por*

$$S(x) = \begin{cases} 2x & si \;\; 0 \leq x < \dfrac{1}{2} \\ 2x - 1 & si \;\; \dfrac{1}{2} \leq x \leq 1 \end{cases} \tag{1.5}$$

es caótico.

Obsérvese que tanto $x = 0$ como $x = 1$ son puntos fijos de $S(x)$. La forma más clara de entender cómo actúa la función (1.5) para el resto de puntos, se obtiene escribiendo la representación binaria de un número $x \in (0, 1)$:

$$x = (0.a_1a_2a_3\dots)_2 = a_1 2^{-1} + a_2 2^{-2} + a_3 2^{-3} + \cdots,$$

donde $a_i \in \{0, 1\}$ para $i \geq 1$. Es conocido que en el sistema decimal un mismo número puede tener dos representaciones. Por ejemplo, se tiene que $1 = 0.9999\cdots = 0.\widehat{9}$ o $0.5 = 0.49999\cdots = 0.4\widehat{9}$, donde $\widehat{\cdot}$ representa la expresión que se repite periódicamente. Lo mismo ocurre en el sistema binario, donde podemos escribir $1/2 = (0.1)_2$ o $1/2 = (0.0\widehat{1})_2$. Para evitar ambigüedades, en números con dos representaciones elegiremos siempre la finita. Así, usaremos $1/2 = (0.1)_2$ y no $(1/2 = 0.0\widehat{1})_2$.

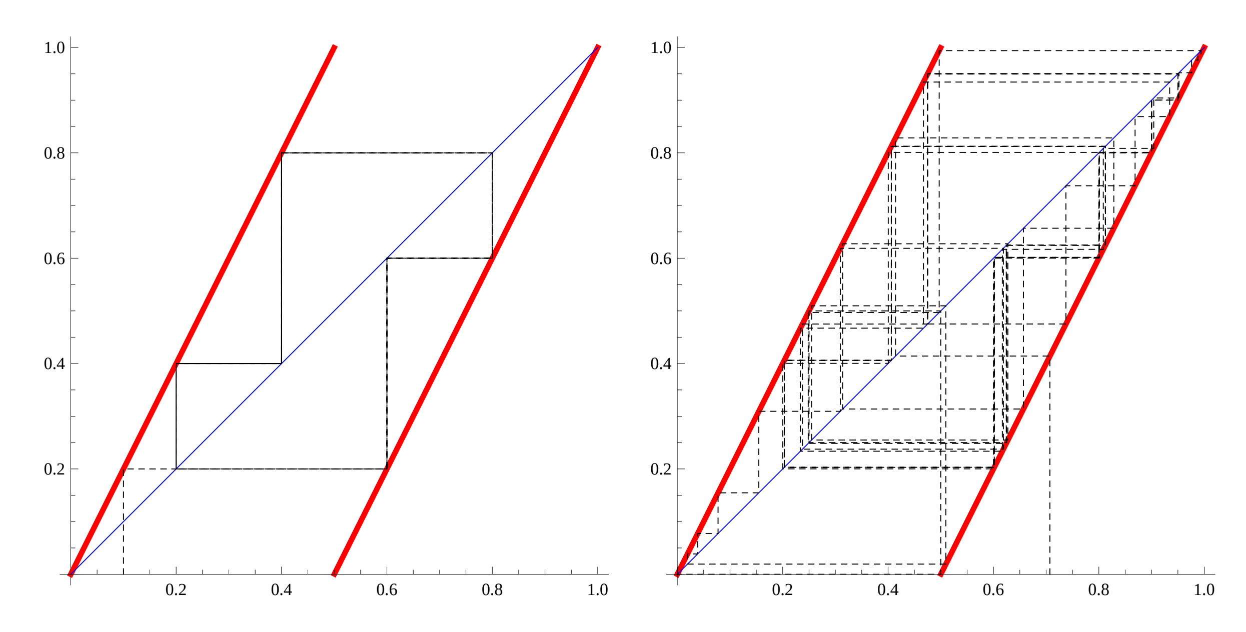


Figura 1.20: Gráfica de la función «diente de sierra» con las órbitas de los puntos $x_0 = 0.1$ y $x_0 = \sqrt{2}/2$. Un simple análisis gráfico parece sugerir un comportamiento periódico en ambos casos. Sin embargo, un análisis más detallado muestra que esto no es así.

Notemos que si $x = (0.a_1a_2a_3\dots)_2$, entonces $S(x) = (0.a_2a_3\dots)_2$. La transformación ha consistido en desplazar el punto binario una unidad hacia la derecha y eliminar el dígito que queda en la parte entera (a la izquierda del punto binario). Por este motivo, a esta transformación se le llama también *operador desplazamiento* (*shift operator*).

Como todo número racional tiene una expresión en binario que es finita o periódica, los números racionales son puntos periódicos de $S(x)$. Por ejemplo, además de los dos únicos puntos fijos, 0 y 1, tenemos un 2-ciclo formado por: $1/3 = (0.\widehat{01})_2$ y $2/3 = (0.\widehat{10})_2$. Hay dos 3-ciclos, uno que comienza por el punto $1/7 = (0.\widehat{001})_2$ y otro que comienza por el punto $3/7 = (0.\widehat{011})_2$. En concreto, estos dos 3-ciclos son $\{1/7, 2/7, 4/7\}$ y $\{3/7, 6/7, 5/7\}$, respectivamente.

En general, se puede probar (véase [56] o [115] para un análisis más detallado de estas cuestiones) que los puntos n-periódicos son números racionales de la forma $p/(2^n - 1)$, aunque esta representación no tiene porqué estar escrita de forma irreducible. Por ejemplo, es sencillo comprobar que existen tres 4-ciclos:

$$\left\{\frac{1}{15}, \frac{2}{15}, \frac{4}{15}, \frac{8}{15}\right\}, \quad \left\{\frac{3}{15}, \frac{6}{15}, \frac{12}{15}, \frac{9}{15}\right\}, \quad \left\{\frac{7}{15}, \frac{14}{15}, \frac{13}{15}, \frac{11}{15}\right\}.$$

Todos son de la forma $p/(2^4 - 1)$. Sin embargo, los posibles numeradores p de estas fracciones no recorren el conjunto $\{1, \dots, 14\}$. Esto es debido a que algunas fracciones de la forma $p/15$ ya habían aparecido en ciclos más cortos. En concreto $5/15 = 1/3$ y $10/15 = 2/3$ eran los componentes del 2-ciclo existente.

Llegados a este punto, queremos hacer notar que si $x \in (0,1)$ es un punto con representación binaria infinita y no periódica, entonces la órbita $S^n(x)$ no puede ser periódica. El punto $x = \sqrt{2}/2$ considerado en la segunda gráfica de la figura 1.20 es de estas características, por tanto, su órbita no puede ser periódica, tal y como sugería la representación gráfica.

El hecho conocido de que los números racionales (que son los puntos periódicos de $S(x)$) son densos en los reales ya nos permite aventurar que el sistema dinámico $([0,1], S)$ es caótico. Para demostrarlo con más rigor, veamos que dados dos conjuntos abiertos U y V cualesquiera de $[0,1]$, existe una órbita periódica que visita ambos. En efecto, sean $x = (0.a_1a_2a_3\dots)_2 \in U$ e $y = (0.b_1b_2b_3\dots)_2 \in V$. Como U y V son abiertos, existe $n \in \mathbb{N}$, lo suficientemente grande, tal que el punto

$$z = (0.a_1a_2\dots a_nb_1b_2\dots b_na_1a_2\dots a_nb_1b_2\dots b_n\dots)_2 \in U$$

ya que $|z - x| \leq 2^{-n}$. Pero por otra parte, la órbita de z también tiene puntos de V ya que

$$S^n(z) = (0.b_1b_2\dots b_na_1a_2\dots a_nb_1b_2\dots b_na_1a_2\dots a_n\dots)_2 \in V$$

y entonces $|S^n(z) - y| \leq 2^{-n}$.

Por último, notemos que z es periódico ya que

$$S^{2n}(z) = (0.\overline{a_1a_2\dots a_nb_1b_2\dots b_n})_2 = z.$$

En consecuencia, aplicando el teorema 1.8, el sistema dinámico asociado a la función «diente de sierra» definida en (1.5) es caótico.

Ejemplo 1.17. *El sistema dinámico (X,T), donde $X = [0,1]$ y T es la función conocida como «tienda de campaña», definida por*

$$T(x) = \begin{cases} 2x & si \quad 0 \leq x < \dfrac{1}{2} \\ 2(1-x) & si \quad \dfrac{1}{2} \leq x \leq 1 \end{cases} \tag{1.6}$$

es caótico.

El comportamiento de la función «tienda de campaña» es, en cierto modo, similar al de la función «diente de sierra» (1.5) ya que las órbitas de los puntos racionales son periódicas. De hecho, para estudiar el comportamiento de la función (1.6) nos podemos apoyar en el ejemplo anterior. No es complicado probar por inducción que

$$T^{k+1}(x) = T(S^k(x)), \quad x \in [0,1], \quad k \geq 1.$$

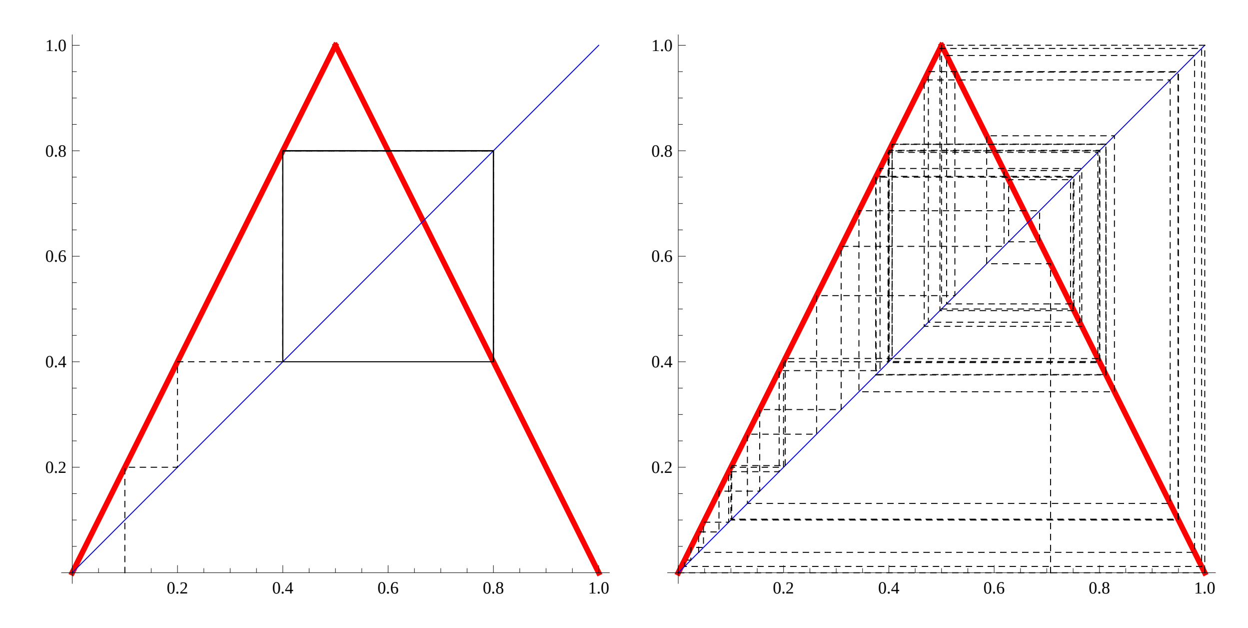


Figura 1.21: Gráfica de la función «tienda de campaña» con las órbitas de los puntos $x_0 = 0.1$ y $x_0 = \sqrt{2}/2$. Al igual que en la figura 1.20, sólo las órbitas de puntos racionales tienen un comportamiento periódico. Los puntos irracionales, como $x_0 = \sqrt{2}/2$, tienen órbitas no periódicas.

Para demostrarlo, aplicamos de nuevo el teorema 1.8. Sean $x = (0.a_1a_2a_3\ldots)_2 \in U$ e $y = (0.b_1b_2b_3\ldots)_2 \in V$ dos puntos de dos conjuntos abiertos cualesquiera U y V de $[0,1]$. Entonces, existe $n \in \mathbb{N}$, lo suficientemente grande, tal que el punto

$$z = (0.a_1a_2\ldots a_n0b_1b_2\ldots b_n0a_1a_2\ldots a_n0b_1b_2\ldots b_n0\ldots)_2 \in U$$

ya que $|z - x| \leq 2^{-n}$. Pero por otra parte, la órbita de z también tiene puntos de V ya que

$$T^{n+1}(z) = T(S^n(z)) = T(0.0b_1b_2\ldots b_n0a_1a_2\ldots a_n0)_2$$

$$= (0.b_1b_2\ldots b_n0a_1a_2\ldots a_n0)_2 \in V$$

puesto que $|T^{n+1}(z) - y| \leq 2^{-n}$.

Además, z es periódico ya que

$$T^{2n+2}(z) = T(S^{2n+1}(z)) = T(0.0a_1a_2\ldots a_n0b_1b_2\ldots b_n0)$$

$$= 0.a_1a_2\ldots a_n0b_1b_2\ldots b_n0 = z.$$

En definitiva, hemos probado que el sistema dinámico asociado a la función «tienda de campaña» definida en (1.6) es caótico.

Ejemplo 1.18. *El sistema dinámico (X, ℓ_4), donde $X = [0,1]$ y ℓ_4 es la función logística de parámetro $\lambda = 4$, $\ell_4(x) = 4x(1-x)$, es caótico.*

Para demostrarlo, nos apoyamos de nuevo el teorema 1.8 y en el hecho de que la función «tienda de campaña» $T(x)$, definida en el ejemplo anterior, es caótica. Además, notemos que T y ℓ_4 son topológicamente conjugadas mediante el homeomorfismo $h(x) = \text{sen}^2(\pi x/2)$. En efecto, se tiene que

$$h \circ T(x) = \ell_4 \circ h(x), \quad x \in [0, 1].$$

Es más, $h \circ T^k(x) = \ell_4^k \circ h(x)$ para todo $k \geq 1$. Por lo tanto, si x_0 es un punto n-periódico de T, entonces $h(x_0)$ es un punto n-periódico de ℓ_4. Teniendo esto en cuenta, sean ahora U y V dos abiertos cualesquiera de X. Denotamos $U' = h^{-1}(U)$ y $V' = h^{-1}(V)$ a los correspondientes abiertos obtenidos a partir de la inversa del homeomorfismo h. Como la función T es caótica, existe $x_0 \in U'$ tal que $T^n(x_0) = x_0$ y además $T^k(x_0) \in V'$. Pero entonces $h(x_0) \in U$ cumple $\ell_4^n(x_0) = x_0$ y $\ell^k(x_0) \in V$. Aplicando el teorema 1.8 se deduce que ℓ_4 es una aplicación caótica en il intervalo $= [0, 1]$.

Ejemplo 1.19. *La función de iteración del método de Newton*

$$N_f(x) = x - \frac{f(x)}{f'(x)}$$

para resolver una ecuación no lineal $f(x) = 0$ puede presentar un comportamiento caótico, para una particular elección de $f(x)$.

Por ejemplo, no hay más que aplicar dicho método a la función

$$f_\mu(x) = \left(\frac{\mu(x-1)+1}{x}\right)^{1/(\mu-1)}$$

donde $\mu > 1$ y $x \neq 0$.

Unos sencillos cálculos nos muestran que

$$N_{f_\mu}(x) = \mu x(1-x),$$

que no es otra que la función logística estudiada en la sección 1.4 y que es la representante obligada cuando se habla de aplicaciones caóticas.

Capítulo 2

El método de Newton

2.1. Introducción

La resolución de ecuaciones no lineales es uno de los problemas matemáticos que más frecuentemente aparece en diversas disciplinas científicas. Así, con la notación

$$f(x) = 0 \tag{2.1}$$

englobamos el problema de encontrar una incógnita x, que puede ser un número real o complejo, un vector, una función, entre otros, a partir de los datos que nos proporciona la función f, que puede ser, por ejemplo, una función escalar, un sistema de ecuaciones, una ecuación diferencial, una ecuación integral, etc.

Incluso en el caso de que f sea una función real de variable real es bien conocido que, en general, no es posible resolver de forma exacta una ecuación no lineal. Es por ello que se recurre a técnicas iterativas para obtener aproximaciones de la solución.

Sin duda, dentro de estas técnicas iterativas, el método de Newton es el procedimiento más estudiado y empleado en la práctica. Así, con el objetivo de aproximar una solución α de una ecuación no lineal (2.1), el método de Newton consiste en construir, a partir de una aproximación inicial x_0 de α, una sucesión de la forma

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0. \tag{2.2}$$

En condiciones adecuadas, la sucesión anterior converge a la solución buscada α.

Es costumbre extendida entre los investigadores el bautizar sus descubrimientos con su propio nombre o con el de un personaje relevante en la materia. En este caso, el nombre de este método está ligado al eminente científico británico Isaac Newton. Sus trabajos de finales del siglo XVII parecen ser el germen del proceso que actualmente lleva su nombre. Sin embargo, tal y como se recogen en numerosos trabajos de índole histórico dedicados a la evolución del método de Newton, éste es fruto de las aportaciones de un gran número de

científicos, tanto anteriores como posteriores a Newton. En la sección 2.2 se presenta de forma detallada la evolución histórica del método (2.2), tal y como se recoge en [46]. Además, las referencias que allí se citan pueden servir al lector para profundizar en este extenso tema.

2.2. Historia del método de Newton

Se ha escrito mucho sobre el origen y evolución de lo que hoy conocemos como método de Newton para resolver ecuaciones no lineales. De hecho, incluso no hay consenso en el nombre del método y muchos autores lo denominan método de Newton-Raphson [31] e, incluso, método de Newton-Raphson-Simpson ([7], [153]). Es más, tal y como podemos encontrar en [84], dos de las características fundamentales del método de Newton empleado en la actualidad son que es iterativo y que incluye el uso de la derivada, aunque ninguna de esas dos peculiaridades fueran observadas por Newton. Las siguientes notas se han extraído del trabajo [46], publicado en la Gaceta de la RSME, donde, entre otras cosas, se presenta el desarrollo histórico del método de Newton.

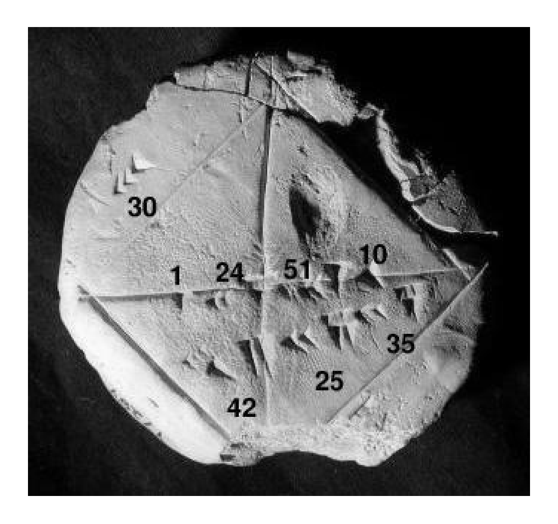


Figura 2.1: Tablilla YBC 7289 de la *Yale Babylonian Collection* (fotografía de Bill Casselman [27]).

Aunque la «paternidad» del método se atribuye a Isaac Newton, la idea de encontrar una cantidad desconocida mediante aproximaciones sucesivas se remonta a muchos siglos atrás. Así, en la Grecia clásica ya se utilizaban técnicas para aproximar números irracionales (sobre todo π) por números racionales. Pero incluso antes, dos mil años antes de Cristo, las civilizaciones mesopotámicas ya conocían técnicas para aproximar la raíz cuadrada de un número. Las referencias al respecto son abundantes. A modo de ejemplo, citamos los textos de Bailey [7], Knill [83] o Neugebauer y Sachs [105]. En particular, en [105, p. 42–43] se pone de manifiesto cómo en la tablilla YBC 7289 (véase la figura 2.1) de la *Yale Babylonian*

Collection [150] aparece un cuadrado de 30 unidades de lado en cuya diagonal están escritos los números $1; 24, 51, 10$ y $42; 25, 35$.[1]

La conversión al sistema decimal de la primera cantidad es $1.4142129629\ldots$, que coincide hasta la quinta cifra decimal con $\sqrt{2} = 1.4142135623\ldots$ Además, la segunda cantidad es el producto de 30 por la primera y es, por tanto, la longitud de la diagonal del cuadrado. Así pues, parece claro que los babilonios conocían un valor aproximado para $\sqrt{2}$ y que lo usaban para sus cálculos.

Otro indicio de que los babilonios sabían cómo aproximar cantidades irracionales aparece en la *tablilla VAT6598* que se conserva en el museo de Berlín y está fechada en 2000–1700 a. C. En ella se plantea, entre otros, el problema de encontrar la diagonal de un rectángulo de altura 40 y lado 10. Con la notación actual, el problema se traduciría en encontrar

$$\sqrt{40^2 + 10^2} = \sqrt{1700}.$$

En la misma tablilla se propone como aproximación el valor $41; 15 = 41 + 15/60$. No se sabe cómo se obtuvo este valor, ni tampoco hay indicios de que se use un proceso iterativo, pero algunos autores [31] destacan el hecho de que esta cantidad coincida con la conocida aproximación para una raíz cuadrada

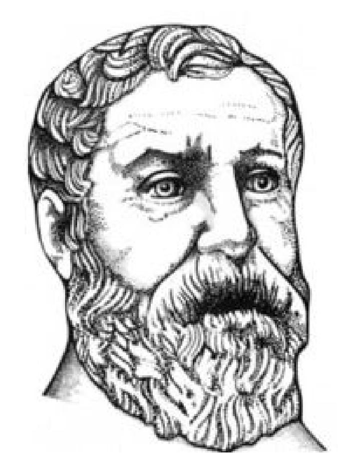
$$\sqrt{h^2 + l^2} \simeq h + \frac{l^2}{2h}$$

para $h = 40$ y $l = 10$.

La aproximación anterior se conoce como *fórmula de Herón* para el cálculo de raíces cuadradas, en la que, partiendo de una aproximación inicial a de $\sqrt{A}$, se propone como nueva aproximación $(a + A/a)/2$. En efecto, para $A = h^2 + l^2$ y $a = h$, la aproximación dada en la tablilla babilónica coincide con la aproximación de Herón. Aunque hay quien atribuye la fórmula de Herón al pitagórico Arquitas de Tarento (428–365 a. C.) o incluso a Arquímedes (282–212 a. C.), [7], donde el método aparece es en el primer tomo de la *Métrica* que Herón publicó en el siglo I. En este texto, descubierto por H. Schöne en 1896 (véase [31] para más detalles) se muestra cómo Herón calculaba el área de un triángulo de lados 7, 8 y 9 unidades, es decir, $\sqrt{720}$. En el mismo, Herón se refiere explícitamente a que una aproximación dada puede ser tomada como punto de partida para obtener mejores aproximaciones. Parece claro, por tanto, que estamos ante la primera referencia de la utilización de un proceso iterativo para resolver un problema.

Ahora bien, ¿fue el método de Herón novedoso en su época? o ¿era una técnica ya conocida y que había sido empleada por civilizaciones anteriores? La respuesta queda en el aire,

[1] Los babilonios usaban un sistema de numeración cuneiforme de base sexadecimal. En la actualidad, los expertos en el tema escriben los números babilónicos usando una mezcla de nuestra notación en base 10 y su notación en base 60. El equivalente babilónico de la coma decimal se denota con un punto y coma. El resto de los dígitos se separan por comas. Así, el número $5, 51, 13; 2, 30$ significa $5 \times 60^2 + 51 \times 60 + 13 + 2 \times 1/60 + 30 \times 1/60^2 \simeq 21073.0416$.



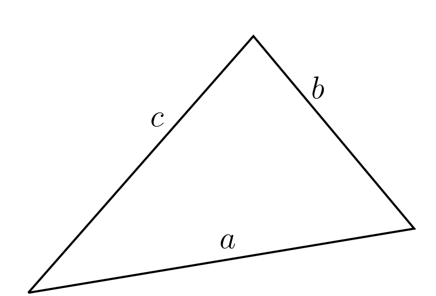


Figura 2.2: Herón de Alejandría (10–70 d. C. aproximadamente) y su fórmula para calcular el área A de un triángulo, conocidos sus lados a, b y c: $A = \sqrt{s(s-a)(s-b)(s-c)}$, con $s = (a+b+c)/2$.

aunque la mayoría de los investigadores en esta parte de la Historia de las Matemáticas parecen inclinarse hacia la segunda opción, ya que hay constancia del uso de textos babilónicos por parte de matemáticos y astrónomos contemporáneos con Herón. Por ejemplo, Claudio Ptolomeo (100–170 d. C.) cita en su *Almagesto* datos astronómicos de la época del rey asirio Nabonassar (747 a. C.).

A partir de la fórmula de Herón, las técnicas para calcular la raíz cuadrada de un número (y, en general, las raíces n-ésimas) se fueron transmitiendo y/o redescubriendo a través de los siglos y de las civilizaciones hasta el siglo XVII. Aunque no hay muchas evidencias escritas de lo que ocurrió durante este largo período de tiempo, sí que, a modo de ejemplo, se pueden encontrar algunas referencias sobre métodos para el cálculo de raíces n-ésimas [31].



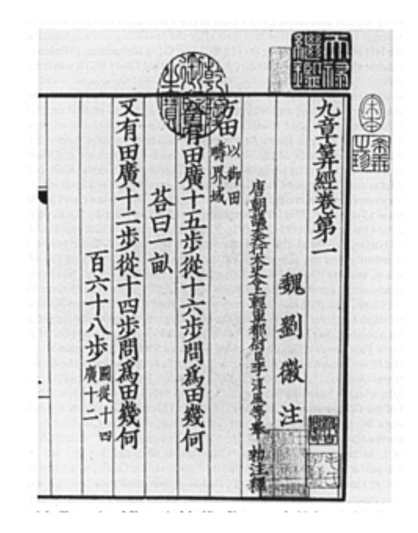

九章算經卷第一
魏 劉徽 注
唐朝議大夫行太史令上輕車都尉臣李淳風等奉 勅注釋
方田 以御田疇界域
今有田廣十五步從十六步問爲田幾何
荅曰一畝
又有田廣十二步從十四步問爲田幾何
百六十八步 圖從十四廣十二

Figura 2.3: Sello con la figura de Liu Hui (220–280 d. C. aproximadamente), autor de una versión comentada del *Jiuzhang suanshu*, una de cuyas páginas se muestra a su derecha.

Podemos citar, por ejemplo, el libro chino de Matemáticas por excelencia, el *Jiuzhang*

suanshu, que se traduce por *Nueve capítulos del arte matemático*. Se conoce una versión del siglo III, con comentarios de Liu Hui (220–280 d. C. aproximadamente), que contiene una colección de problemas que requieren el cálculo de raíces cuadradas y cúbicas [107]. Posteriormente, en el siglo IV, Teón de Alejandría (335–405 d. C. aprox.), padre de Hypatia, desarrolló un método totalmente geométrico para el cálculo aproximado de raíces cuadradas. En los trabajos del matemático persa Sharaf al-Din al-Tusi (1135–1213) se encuentran las soluciones, tanto algebraicas como numéricas, de algunas ecuaciones cúbicas. Además, parece ser que al-Tusi fue el primero en calcular la derivada de un polinomio de tercer grado.

En el trabajo *Raf al-Hijab* del matemático árabe Al-Marrakushi ibn Al-Banna (1256–1321), que podríamos traducir por *Que levanta el velo*, se indica cómo calcular raíces cuadradas usando series y fracciones continuas. Parece ser que Al-Banna fue un gran recopilador de los conocimientos matemáticos de su época, y que en sus escritos nos muestra su versión de los trabajos de matemáticos árabes anteriores.

El problema de encontrar la raíz n-ésima de un número fue evolucionando hacia el problema más general de encontrar las raíces de una ecuación polinómica e, incluso, de una ecuación trascendental, como por ejemplo la conocida ecuación de Kepler, que formalmente se puede escribir como

$$f(x) = x + a \operatorname{sen} x - b$$

y de la que hablaremos con más detalle en el ejemplo 2.2.

En esta época son numerosos los intentos para encontrar de manera algebraica las soluciones de ecuaciones polinómicas. Como anécdota, se puede destacar la pugna que mantuvieron algunos matemáticos del Renacimiento italiano para resolver las ecuaciones de tercer y cuarto grado. Por entonces era habitual que los matemáticos participaran en competiciones públicas para demostrar su ingenio, además de para poder ganar dinero en las apuestas que se suscitaban (véase [140] para más detalles). En 1535 tuvo lugar una de estas competiciones en las que participaron los matemáticos Antonio Fiore y Niccolo Fontana, apodado *Tartaglia*, el «tartamudo». En este tipo de competiciones, cada participante proponía problemas a su oponente, y el que más resolviera era el ganador. Los problemas que se plantearon Fiore y Tartaglia hacían referencia a la resolución de ecuaciones cúbicas. Por entonces, las ecuaciones cúbicas estaban clasificadas en tres tipos, de los cuales Fiore sólo sabía resolver uno y Tartaglia había aprendido a resolver los tres. Evidentemente, Tartaglia fue el vencedor de la competición, pues propuso a su oponente ecuaciones de los tipos que no sabía resolver.

La noticia del duelo Fiore-Tartaglia llegó a oídos de otro matemático de la época, Gerolamo Cardano, que estaba escribiendo un texto sobre álgebra. Cardano quiso añadir a su libro los métodos para resolver ecuaciones cúbicas. Para ello le pidió a Tartaglia que se los comunicara. Pero Tartaglia era algo reacio a ello, no en vano sus técnicas eran una forma de ganarse la vida. Finalmente Tartaglia accedió y transmitió a Cardano sus conocimientos, con la condición de que éste no los publicara. Pero años más tarde, en 1545, Cardano publicó su




Figura 2.4: Gerolamo Cardano (1501–1576) y Niccolo Fontana (1500–1557), alias *Tartaglia*, dos de los protagonistas de la intriga renacentista sobre el descubrimiento de las soluciones de una ecuación cúbica.

Ars Magna, en la cual aparecían las soluciones de las ecuaciones de tercer y cuarto grado. La polémica estaba servida y Tartaglia acusó a Cardano de plagio, sin que el hecho de que Cardano hubiese dado todo el crédito del descubrimiento a Tartaglia fuese un atenuante. Aún así, Cardano supo defenderse y justificó la ruptura de su pacto con Tartaglia. En efecto, con anterioridad a la publicación de su obra *Ars Magna*, llegaron a manos de Cardano los escritos inéditos de otro matemático, fechados con anterioridad a su encuentro con Tartaglia y en los que, de forma independiente, se resolvían los tres tipos de ecuaciones cúbicas. Este matemático era Scipione del Ferro, a la sazón maestro de Antonio Fiore. Por lo tanto, Cardano argumentó que había publicado el método del del Ferro, no el de Tartaglia. Además, Cardano tenía otra buena razón para romper su pacto con Tartaglia: Ludovico Ferrari, discipulo de Cardano, había descubierto un método para resolver las ecuaciones cuárticas. Este método se reducía a la resolución de una ecuación cúbica asociada. Por lo tanto, para publicar el descubrimiento de Ferrari, Cardano debía hablar inevitablemente del método de Tartaglia.

Tartaglia no entendió así las cosas y llegó a insultar públicamente a Cardano, tanto personal como profesionalmente. El conflicto terminó con Ferrari retando a un debate público a Tartaglia, para defender a su maestro. Ferrari venció con solvencia y Tartaglia se tuvo que retirar de la escena científica. Como consecuencia, las fórmulas de Tartaglia se conocen actualmente como fórmulas de Cardano.

A partir del siglo XV, el problema se fue bifurcando en varias líneas (resolución algebraica de ecuaciones polinómicas, resolución aproximada usando iteraciones de punto fijo, aproximaciones mediante fracciones continuas, etc.). Para un análisis detallado del desarrollo histórico de estos problemas debemos remitir al lector interesado a alguno de los textos especializados, como [31] o [153].

Centrándonos en el «nacimiento» del método de Newton, podemos destacar el antecedente del matemático francés François Viète (1540–1603), quien desarrolló un ambicioso proyecto



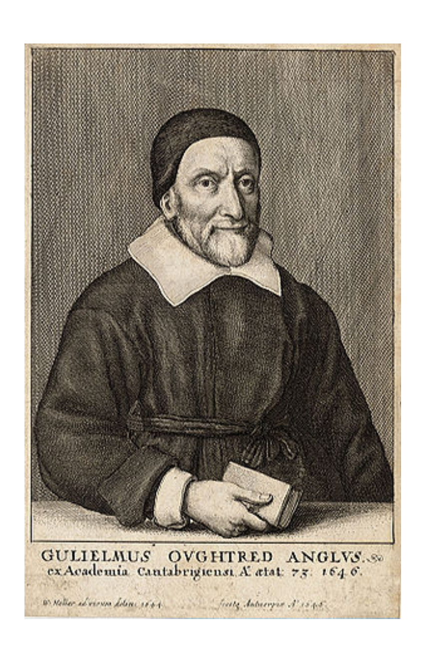


Figura 2.5: François Viète (1540–1603) y William Oughtred (1574–1660) se encuentran entre los antecesores del método de Newton. Para muchos historiadores, Viète fue quien inspiró a Newton para desarrollar su método.

para buscar, de forma genérica, las soluciones positivas de ecuaciones polinómicas de grados 2 a 6. Viète fue el primero en representar los parámetros de una ecuación con letras, no sólo las incógnitas. El método empleado por Viète («logística especiosa» o «arte del cálculo sobre símbolos») estaba inspirado en la tradición geométrica griega. El método de Viète, escrito en un lenguaje arcaico y con notaciones engorrosas, no tuvo continuación y pronto pasó al ostracismo, apartado por la geometría cartesiana. Sin embargo, Viète fue el primero en darse cuenta de la relación existente entre las raíces y los coeficientes de un polinomio, e intentó usar el álgebra para la resolución de estos problemas.

El procedimiento expuesto por Viète fue simplificado por otros matemáticos, entre los que podemos destacar al inglés William Oughtred (1574–1660). En efecto, su obra *Clavis Mathematicae*, publicada en 1631, fue considerada como un texto de referencia en la época y fue empleada por científicos como Wallis o Newton. En ella Oughtred apuesta por un estilo más ligero a la hora de escribir matemáticas, innovando en las notaciones que empleaba, en contraposición con el estilo más recargado empleado por Viète.

Parece ser que el trabajo de Viète fue el que inspiró a Isaac Newton (1643–1727) para desarrollar su método para resolver ecuaciones. La primera referencia escrita sobre el método de Newton se encuentra en *De analysi per aequationes numero terminorum infinitas*, en una carta escrita a sus colegas Barrow y Collins en 1669, pero que no fue publicada hasta 1711. A los dos años de escribir la citada carta, en 1671, Newton desarrollaba su método en *De metodis fluxionum et serierum infinitarum*. De nuevo, esta obra tardó en publicarse y no es hasta 1736 cuando se publica una traducción de la misma bajo el título de *Method of Fluxions*. Como se pone de manifiesto en [25], la primera versión impresa del método de aproximación

de Newton, aparece en el capítulo 94 del *Algebra* de J. Wallis, publicada en 1685.

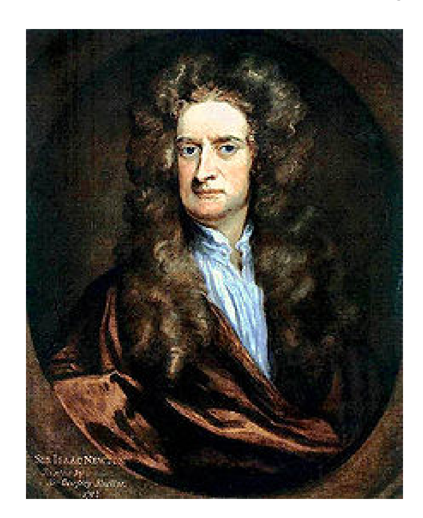




Figura 2.6: Isaac Newton (1643–1727), el «padre» del conocido método para resolver ecuaciones no lineales que lleva su nombre. A la derecha, su estatua en la capilla del Trinity College de Cambridge realizada en 1755 por Louis Francois Roubiliac.

Para hacernos una idea de cómo trabajaba Newton, podemos ilustrar su método con el mismo ejemplo que él consideró, la ecuación $x^3 - 2x - 5 = 0$. Newton argumentaba de la siguiente manera:

> Por tanteo, se ve que la solución está cerca de 2. Haciendo $x = 2+\varepsilon$ y sustituyendo en la ecuación se obtiene:
>
> $$\varepsilon^3 + 6\varepsilon^2 + 10\varepsilon - 1 = 0. \tag{2.3}$$
>
> Ignorando los términos $\varepsilon^3 + 6\varepsilon^2$ con el pretexto de que ε es pequeño, se llega a que $10\varepsilon - 1 \simeq 0$ ó $\varepsilon = 0.1$. Entonces, $x = 2.1$ es una aproximación de la solución mejor que la inicial.
>
> Haciendo ahora $\varepsilon = 0.1 + \nu$ y sustituyendo en (2.3) se sigue que
>
> $$\nu^3 + 6.3\nu^2 + 11.23\nu + 0.061 = 0.$$
>
> Ignorando de nuevo los términos en ν de grado mayor o igual que dos, se llega a que $\nu \simeq -0.054$ y, por tanto, $x = 2.046$ es una aproximación que mejora las anteriores. Newton indicaba que el proceso se puede repetir las veces que sean necesarias.

Como vemos, la idea de Newton consiste en añadir un término corrector a una aproximación inicial dada. Para obtener esta aproximación, lo que hace es truncar el binomio de Newton en el segundo término, en expresiones del tipo

$$(a + \varepsilon)^n \simeq a^n + na^{n-1}\varepsilon.$$

De esta manera, para obtener el valor aproximado de ε, simplemente hay que resolver una ecuación lineal.

Escribiendo el problema con la notación actual y llamando $p(x) = x^3 - 2x - 5$, tenemos que la nueva aproximación es

$$2 - \frac{p(2)}{p'(2)} = 2 + \frac{1}{10} = 2.1,$$

que se corresponde con la conocida formulación del método de Newton (2.2) cuando $f(x)$ es el polinomio $p(x)$ anterior. Ahora bien, no se tiene constancia de que Newton usara el cálculo diferencial ni de que expresara el proceso como un método iterativo en el sentido de que una aproximación pueda ser considerada como punto de partida de la siguiente. Además, Newton usaba «su método» sólo para resolver ecuaciones polinómicas. El propio Newton era consciente de que su método podía fallar y lo justifica desde el principio de su obra *De analysi per aequationes numero terminorum infinitas*, donde encontramos la siguiente cita donde aclara que lo que ha escrito está brevemente explicado más que demostrado con toda diligencia y que lo que muestra es un método general que cavilará un día para medir la cantidad de las curvas mediante una serie de términos infinitos:

> «Methodum generalem, quam de curvarum quantitate per Infinitam terminorum Seriem mensurada, olim excogitaveram, in sequentibus breviter explicatam potius quam accurate demonstratam habes.»

Según algunos historiadores como [25] o [128] ésta es una prueba de que en esa época primaba más el interés de los matemáticos por el descubrimiento que por la justificación rigurosa de lo descubierto. Por todo ello, podemos concluir que la idea que Newton tenía de su método dista bastante de la que tenemos hoy en día.

La idea de iteración se atribuye (véanse, por ejemplo, [31] y [153]) a Joseph Raphson[2] (1648–1715), quien además simplifica el aspecto operacional de la técnica de Newton. En 1690 publica un tratado, *Analysis Aequationum Universalis*, en el que se dan fórmulas explícitas para el término corrector para algunos casos particulares de ecuaciones. En concreto, calcula los términos correctores para las ecuaciones $x^3 - r = 0$ y $x^3 - px - q = 0$ que son, respectivamente,

$$\frac{r - x_0^3}{3x_0^2} \quad \text{y} \quad \frac{q + px_0 - x_0^3}{3x_0^2 - p},$$

siendo x_0 la aproximación inicial.

Notemos que Raphson publicó su obra 46 años antes que el «Método de las fluxiones» de Newton. Esta publicación no estuvo exenta de unas pequeñas dosis de polémica. En efecto,

[2]Nos hubiera gustado incluir una fotografía de Raphson, pero no hemos encontrado ninguna disponible en la red, como se indica en la página que Wikipedia dedica a este científico, `http://en.wikipedia.org/wiki/Joseph_Raphson`

como se menciona en [128], el propio Raphson, en la presentación de su método a la Royal Society en 1690, es el primero en reconocer que el método de Newton ya era conocido en los ambientes científicos de la época y que su método era una versión mejorada, eso sí. Así, para no pecar de plagio, Raphson cita en su prefacio a Newton, entre otros matemáticos. Sin embargo, en la edición de 1697, Raphson solo se refiere a Viète como el antecesor del método. Más adelante cita a Harriot y Oughtred, pero solamente cita a Newton en el apéndice final y en relación al teorema del binomio.

La contribución de Raphson ha sido tenida en cuenta históricamente, no en vano muchos autores denominan el proceso como método de Newton-Raphson, sobre todo a partir del trabajo de Cajori [25]. Sin embargo, en los trabajos de Raphson no se aprecia la conexión existente entre el término corrector, la función que define la ecuación, y su derivada.

La incorporación del cálculo diferencial se debe a Thomas Simpson (1710–1761). Como se puede ver en [84] o [153], Simpson, en su obra *Essays on Mathematics*, publicada en 1740, fue quien estableció el método tal y como lo conocemos actualmente, salvo aspectos notacionales (Simpson explicaba de forma retórica cómo obtener las aproximaciones sucesivas). Además, Simpson extendió el proceso a funciones cualesquiera, no solamente polinomios.



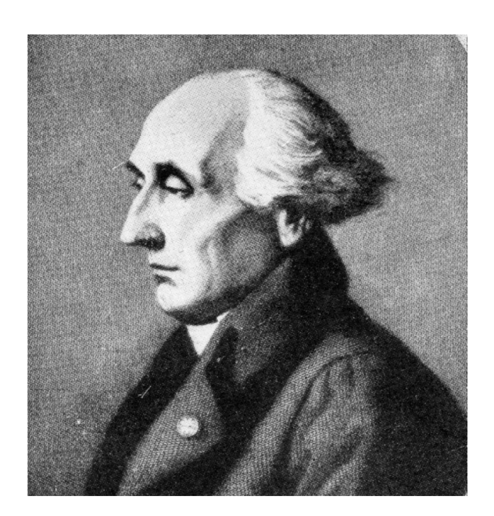


Figura 2.7: A la izquierda, Thomas Simpson (1710–1761), el «gran olvidado» en cuanto a sus aportaciones a lo que hoy conocemos como método de Newton. A la derecha, Joseph Louis Lagrange (1736–1813), el primero en plantear interrogantes sobre la convergencia del método de Newton.

Con motivo de ciertas observaciones a propósito de la utilización de series infinitas, Newton parece estar preocupado por el concepto de convergencia, pero no aporta ninguna solución a este problema. La primera vez que aparece la discusión de la convergencia del método de Newton es en 1768, en el *Traité de la résolution des équations en general* del francés Jean Raymond Mourraille (1720–1808). A pesar de contener ideas novedosas, la mayor parte del trabajo Mourraille pasó inadvertido.

Contrariamente a Newton y Raphson, Mourraille hace hincapié en el aspecto geométrico del método de Newton, justificando que éste sea también conocido como método de la tangente. Mourraille utiliza la representación geométrica para explicar el comportamiento de la sucesión iterativa producida por el algoritmo de Newton. Además, Mourraille observa por primera vez que, dependiendo del punto de salida elegido, la sucesión generada por el método puede converger a alguna de las raíces de la ecuación, oscilar, tender a infinito, o acercarse a un límite que no es solución de la ecuación. Finalmente, Mourraille también muestra que la convergencia puede ser más o menos rápida, pero solamente lo indica en forma cuantitativa.

Posteriormente Joseph-Louis Lagrange (1736–1813), publica su *Traité de la résolution des équations numériques de tous les degrés* en 1808 [54]. Esta obra es un compendio de trabajos sobre la teoría de ecuaciones algebraicas en el que aparecen, entre otros, métodos para separar las raíces reales de una ecuación algebraica o métodos de aproximación de raíces mediante fracciones continuas (para más información sobre la obra de Lagrange, se pueden consultar las referencias [90], [128]). En ella Lagrange afirma que el método atribuido a Newton es el que se emplea habitualmente para resolver ecuaciones numéricas. Ahora bien, advierte que este método sólo se puede usar para ecuaciones que están ya «casi resueltas» en el sentido de que para aplicarlo se necesita una buena aproximación de la solución. Además, plantea dudas sobre la exactitud de cada nueva iteración y observa que el método puede tener problemas para el caso de raíces múltiples o muy próximas entre sí. También se preocupó por indicar los procedimientos para encontrar las soluciones complejas de una ecuación algebraica.

Lagrange terció en la polémica sobre la asignación del método a Newton, Raphson o Simpson. En concreto afirma que es un hecho constatable que el método de aproximación de Newton había aparecido en la versión inglesa del *Algebra* de Wallis, publicada en 1685 y en su posterior edición en latín de 1693. Manifiesta su sorpresa por el hecho de que Raphson no mencione a Newton en su obra de 1695, lo cual hace decir a Lagrange que Raphson llegó a pensar que su método era diferente al de Newton. Sin embargo, el propio Lagrange afirma que ambos métodos son el mismo, aunque presentados de forma diferente.

Jean Baptiste Joseph Fourier (1768–1830) fue el primero en analizar la velocidad de convergencia del método de Newton en una nota titulada *Question d'analyse algébraique* (1818) [54]. En este trabajo, Fourier expresa el método con la notación actual (2.2) y lo bautiza como *la méthode newtonienne*, haciendo referencia explícita a las obras de Newton, Raphson y Lagrange. Quizás, Fourier es el «causante» de la falta de reconocimiento para el trabajo de Simpson.

El siguiente matemático importante en estudiar el método de Newton fue Augustin Louis Cauchy (1789–1857), quien estudió este tema desde 1821, pero no dio una formulación satisfactoria hasta la publicación de las *Leçons sur le Calcul différentiel* en 1829 [54]. Cauchy da condiciones, en términos de las derivadas f' y f'', para asegurar que el método de Newton es convergente a una solución α de la ecuación (2.1) para todo punto de partida x_0 perteneciente a un intervalo determinado. Lo que Cauchy estaba buscando son, por tanto, resultados de




Figura 2.8: Jean Baptiste Joseph Fourier (1768–1830), quien «bautiza» a (2.2) como método de Newton y Augustin Louis Cauchy (1789–1857), los primeros en establecer condiciones para garantizar la convergencia del método de Newton.

convergencia global para el método de Newton; es decir, caracterizar los intervalos $I_\alpha \subseteq \mathbb{R}$ para los que

$$\lim_{n\to\infty} x_n = \alpha, \quad \text{con } x_0 \in I_\alpha.$$

Aunque la mayoría del trabajo de Cauchy se centra en el campo real, al final del mismo dedica un apartado al estudio de raíces complejas.

Los primeros avances en el estudio del método de Newton para resolver ecuaciones complejas se dan con los trabajos de E. Schröder y A. Cayley en 1870 y 1879 respectivamente. De hecho, el que se conoce como «problema de Cayley» consiste en caracterizar las regiones S_α del plano complejo para las cuales el método de Newton converge a la raíz α si $x_0 \in S_\alpha$. Este problema, inicialmente resuelto por Cayley para polinomios de segundo grado, encerraba ciertas sorpresas para polinomios de tercer grado. En palabras del propio Cayley, en 1890: «Espero poder aplicar esta teoría al caso de una ecuación cúbica, pero los cálculos son mucho más difíciles». Cuarenta años más tarde, los trabajos de Gaston M. Julia (1918) y Pierre J. L. Fatou (1920) revelaron que el problema al que se enfrentaba Cayley era prácticamente inabordable con los conocimientos y técnicas de su época. Como se indica en el capítulo 4, el estudio de las dinámicas del método de Newton para funciones de variable compleja no es en absoluto trivial y requiere de conocimientos y resultados que no se desarrollaron hasta bien entrado el siglo XX.

El hecho de pensar en el método de Newton para encontrar las raíces de una función de variable compleja nos lleva, de forma natural, a pensar también en el método de Newton para encontrar las raíces de una función vectorial de dos variables $f : \mathbb{R}^2 \to \mathbb{R}^2$. En este caso, el método de Newton admite unas interesantes interpretaciones geométricas, como puede verse en [1] y [152]. Volviendo al punto de vista histórico, fue el propio Simpson quien estudió por primera vez un sistema de dos ecuaciones transcendentales. En el sexto de sus ensayos, escritos en 1740, Simpson describe la técnica para resolver sistemas no lineales de dos ecuaciones con

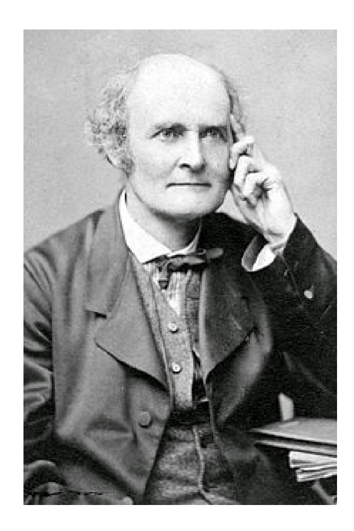




Figura 2.9: Arthur Cayley (1821–1895) y Ernst Schröder (1841–1902), precursores en la extensión del método de Newton al plano complejo.

dos incógnitas usando lo que hoy en día llamamos el método de Newton para sistemas

$$F(x) = 0 \text{ con } F : \mathbb{R}^d \to \mathbb{R}^d.$$

De hecho (véase [153] para más detalles), Simpson indica cómo calcular los elementos de la matriz jacobiana $F'(x)$ para el caso $n = 2$ y la solución del correspondiente sistema de ecuaciones $F'(x)(x_{n+1} - x_n) = -F(x_n)$ usando la regla de Cramer. Simpson ilustra su técnica con tres ejemplos. El primero de ellos, con la notación actual, es:

$$\begin{cases} y + \sqrt{y^2 - x^2} = 10, \\ x + \sqrt{y^2 + x} = 12. \end{cases}$$

A pesar del carácter innovador del trabajo de Simpson, éste no tuvo mucha repercusión en su época.

El problema de la extensión del método de Newton al caso multidimensional cayó en el olvido hasta comienzos del siglo XX. En 1916, Henry B. Fine y Albert A. Bennet, ambos profesores de la Universidad de Princeton, publicaron sendos artículos *On Newton's Method of Approximation* y *Newton's method in general analysis* para sistemas de ecuaciones y para ecuaciones funcionales respectivamente. En el trabajo de Fine se da un resultado de convergencia para el método de Newton aplicado a un sistema de ecuaciones en el cual no se asume la existencia de solución, sino que sólo se exigen condiciones sobre el punto de partida. En ese mismo trabajo, Fine afirma que no había encontrado anteriormente resultados de este tipo para funciones vectoriales. En el trabajo de Bennet se justifica el empleo del método de Newton también en el caso de ecuaciones funcionales, extendiendo el trabajo de Fine. Ambos trabajos resultaron muy innovadores en una época en la que el análisis funcional era una rama de las Matemáticas que estaba dando sus primeros pasos. De hecho, sólo unos años más tarde, en 1932, Stefan Banach introdujo la noción de espacios de Banach en su famoso libro *Theorie des opérations linéaires* [9].

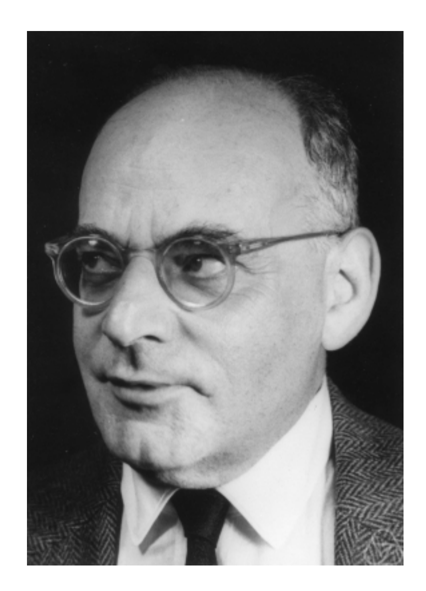




Figura 2.10: Alexander Markowich Ostrowski (1893–1986) y Leonid Vitalievich Kantorovich (1912–1986) extendieron el método de Newton para operadores definidos en espacios de Banach.

A partir de ese momento, varios investigadores muestran interés en la extensión del método de Newton para sistemas de ecuaciones y las publicaciones sobre el tema comienzan a ser numerosas. Pueden consultarse algunas bibliografías específicas sobre el método de Newton ([95], [101]) o buscar en las bases de datos especializadas.

Durante este período, destacamos las publicaciones de Alexander M. Ostrowski [110] y [111], en las que se estudian y comparan distintas condiciones de convergencia dadas previamente por otros autores. Además, se dan estimaciones del error cometido al aplicar el método de Newton para resolver un sistema de ecuaciones no lineales.

Es en este contexto cuando Leonid V. Kantorovich se plantea en [74] la extensión del método de Newton para resolver ecuaciones funcionales definidas entre dos espacios de Banach. Con Kantorovich se inicia el estudio «moderno» del método de Newton con numerosas publicaciones interesantes, tanto desde el punto de vista teórico como práctico. Tal vez, podríamos considerar como las referencias fundamentales en este campo los libros *Functional Analysis in Normed Spaces* [76] y la segunda edición del mismo, *Functional Analysis* [77], que el propio Kantorovich, en colaboración con su discípulo G. P. Akilov, publicaron en 1964 y 1982, respectivamente. En estos textos se recogen los principales resultados sobre el método de Newton y sus aplicaciones obtenidos en los artículos de investigación publicados por el propio Kantorovich entre 1934 y 1957, junto con los conceptos preliminares de Análisis funcional necesarios para su desarrollo. Aunque muchas de las ideas que subyacen bajo la expresión del método de Newton (2.2) definido para ecuaciones escalares se pueden extender a operadores definidos entre espacios de Banach, en el nuevo contexto general hay que tener en cuenta nuevas consideraciones. Por ejemplo, si consideramos una ecuación de la forma $F(x) = 0$, donde $F : X \to Y$ es un operador definido entre dos espacios de Banach X e Y,

la expresión de lo que se conoce como *método de Newton-Kantorovich* queda de la siguiente forma:

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \quad n \geq 0. \tag{2.4}$$

Nótese que, en este caso, $F'(x_n)$ es un operador lineal definido entre X e Y y $F'(x_n)^{-1}$ denota a su inverso. En consecuencia, en cada iteración del método de Newton-Kantorovich hay que invertir un operador lineal o, equivalentemente, resolver la ecuación lineal asociada $F'(x_n)u_n = -F(x_n)$ para definir a continuación $x_{n+1} = x_n + u_n$.

El estudio detallado de la teoría de Newton-Kantorovich escapa del alcance de este texto. No obstante, el lector interesado en tener una visión más general sobre dicha teoría puede consultar algunos artículos de tipo descriptivo como [46], [47] o [120].

A partir de Kantorovich, el número de investigadores que trabajan sobre el método de Newton, sus variantes y aplicaciones comienza a crecer vertiginosamente. Sería prácticamente imposible hacer un seguimiento detallado de las contribuciones de otros autores. Como una pequeña muestra, podemos citar algunos de los autores que han escrito monografías especializadas en el método de Newton y, en general, en procesos iterativos para resolver ecuaciones no lineales: I. K. Argyros [5], J. P. Dedieu [40], J. E. Dennis y R. B. Schnabel [41], A. S. Householder [68], C. T. Kelley [78], J. M. Ortega y W. C. Rheinboldt [109], A. Ostrowski [112], F. A. Potra y V. Pták [121], L. B. Rall [123] y J. F. Traub [146]. Consultando las propias referencias de estos textos, en trabajos recopilatorios como [120] o [151], en sitios web especializados [95] o en bases de datos específicas sobre Matemáticas (Zentralblatt MATH, MathSciNet, Scopus, entre otras) podemos ampliar la larga lista de investigadores con artículos publicados al respecto.

2.3. Construcciones y variantes del método de Newton

En la sección 2.2 ya se ha puesto de manifiesto el gran número de personas que han influido en lo que hoy conocemos como método de Newton para resolver ecuaciones no lineales. Con lo anterior, hemos narrado el inicio de lo que se entiende por método de Newton. En esta sección veremos sus interpretaciones y sus variantes en el estudio de diversos problemas matemáticos. En este listado no precisamos las condiciones que han de cumplirse para que estas construcciones puedan realizarse. Así, supondremos que las funciones que manejaremos son todo lo continuas y derivables que necesiten ser.

- **Construcción geométrica.** El método de de Newton tiene una clara interpretación geométrica basada en la aproximación de la función $f(x)$ por una recta tangente (véase la figura 2.11). Por este motivo, dicho método suele denominarse también método de la tangente. Partiendo de una aproximación inicial x_0 de la solución α, la siguiente aproximación, x_1, se obtiene como la intersección de la recta tangente a la función $f(x)$

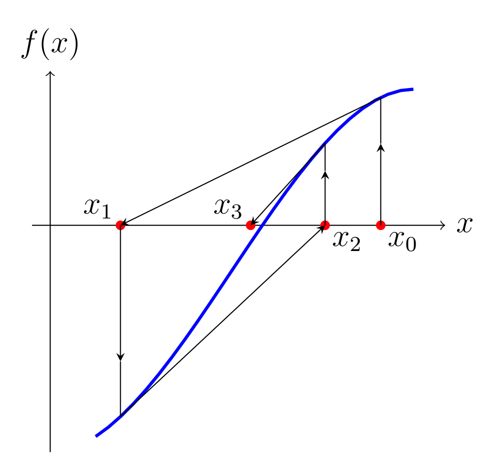


Figura 2.11: Interpretación geométrica del método de Newton.

en el punto x_0 con el eje de abcisas. Así, x_1 es la intersección de las rectas $y - f(x_0) = f'(x_0)(x - x_0)$ e $y = 0$, es decir, $x_1 = x_0 - f(x_0)/f'(x_0)$. Reiterando este procedimiento se obtiene una sucesión de aproximaciones, que converge a la solución y que coincide con la obtenida mediante el método (2.2).

- **Construcción a partir del desarrollo de Taylor.** A partir de x_0 buscamos una aproximación mejor de α que sea de la forma $x_0 + \varepsilon$. Lo ideal sería que $f(x_0 + \varepsilon) = 0$, en cuyo caso se obtendría la solución exacta. Usando el desarrollo de Taylor, llegamos a

$$0 = f(x_0 + \varepsilon) = f(x_0) + f'(x_0)\varepsilon + \frac{f''(x_0)}{2}\varepsilon^2 + \frac{f'''(x_0)}{6}\varepsilon^3 + \cdots \tag{2.5}$$

Truncando (2.5) en el segundo sumando, es decir, linealizando la ecuación, se obtiene:

$$0 \simeq f(x_0) + f'(x_0)\varepsilon \;\Rightarrow\; \varepsilon \simeq -\frac{f(x_0)}{f'(x_0)}.$$

De aquí se deduce la siguiente aproximación de α: $x_0 - f(x_0)/f'(x_0)$. Reiterando el proceso obtenemos el método de Newton (2.2).

- **Construcción a partir de la interpolación racional inversa.** El problema de aproximar una solución de (2.1) se puede transformar en el de aproximar $\phi(0)$, donde $x = \phi(y)$ denota la función inversa de $y = f(x)$. Supongamos que x_0 es una aproximación inicial de la solución α y sea $y_0 = f(x_0)$. Considerando el desarrollo de Taylor de orden uno de la función $\phi(y)$ en torno al punto y_0,

$$\phi(0) \simeq \phi(y_0) - \phi'(y_0)y_0 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

se obtiene una nueva aproximación para $\alpha = \phi(0)$ que coincide, una vez más, con la expresión del método de Newton (2.2).

- **Construcción usando fórmulas de cuadratura.** El método de Newton también se puede construir a partir de fórmulas de integración numérica de tipo interpolatorio. En concreto, si tenemos en cuenta la representación integral

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\,dt$$

y usamos la fórmula de cuadratura de Newton-Cotes de los rectángulos a izquierda,

$$\int_{x_0}^{x} f'(t)\,dt \simeq (x - x_0)f'(x_0),$$

llegamos a que la solución de la ecuación $f(x) = 0$ puede aproximarse por la solución de la ecuación $f(x_0) + (x - x_0)f'(x_0) = 0$, lo que conduce de nuevo al método (2.2).

- **El método de Newton continuo.** El método de Newton (2.2) puede obtenerse también usando el método de Euler explícito para aproximar la solución de la ecuación diferencial

$$x'(t) = -\frac{f(x(t))}{f'(x(t))}, \quad t > 0. \tag{2.6}$$

En efecto, si se considera la condición inicial $x(t_0) = x_0$ y se toma como paso de integración $h = 1$, el método de Euler define una sucesión que coincide con la dada en (2.2). La ecuación (2.6) se conoce como método de Newton continuo y ofrece interesantes interpretaciones desde el punto de vista dinámico. En concreto, si $x(t)$ es una solución de la ecuación (2.6) que cumple $x'(t) \to 0$ cuando $t \to \infty$ se tiene que $x(t)$ «fluye» a una solución de la ecuación (2.1).

Pero no solo son elevados el número de investigadores que han trabajado sobre el método de Newton y las posibles derivaciones de dicho método. Los temas y aplicaciones que se han abordado son numerosos y diferenciados entre sí. La siguiente lista es sólo una pequeña muestra de tópicos relacionados con el método de Newton.

- **Variantes de las condiciones de convergencia del método de Newton-Kantorovich.** De una forma muy esquemática, la teoría de Newton-Kantorovich asegura la convergencia del método (2.4) bajo condiciones que involucran al punto de partida x_0 y al operador F y sus derivadas. Dependiendo de las condiciones exigidas se pueden obtener dominios de puntos de partida x_0 más o menos amplios. Sobre este aspecto se profundizará un poco más en la sección 2.4.

- **Métodos de tipo Newton y métodos con orden de convergencia alto.** Cuando se emplean procesos iterativos para resolver una ecuación no lineal siempre hay que buscar un equilibrio entre el coste computacional que conlleva el proceso y su velocidad de convergencia. En general, los métodos con una mayor velocidad de convergencia llevan asociados un coste computacional alto y viceversa. Bajo unas condiciones bastante

generales, se puede probar que el método de Newton tiene convergencia cuadrática (véase la definición de orden de convergencia en la definición 2.1). Esto quiere decir que el número de cifras significativas en el proceso de aproximar la solución se va multiplicando por dos en cada paso. Aunque el método de Newton se puede considerar como suficientemente rápido, existen otros métodos (Halley, Chebyshev, súper-Halley, etc.) con convergencia cúbica o superior. Todos estos métodos tienen una interpretación geométrica similar a la del método de Newton, como puede verse en [4]. La expresión de las correspondientes funciones de iteración es la siguiente:

1. Método de Halley:
$$x_{n+1} = x_n - \frac{2}{2 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)}, \tag{2.7}$$

2. Método de Chebyshev:
$$x_{n+1} = x_n - \left(1 + \frac{1}{2} L_f(x_n)\right) \frac{f(x_n)}{f'(x_n)}, \tag{2.8}$$

3. Método súper-Halley:
$$x_{n+1} = x_n - \left(1 + \frac{1}{2} \frac{L_f(x_n)}{1 - L_f(x_n)}\right) \frac{f(x_n)}{f'(x_n)}, \tag{2.9}$$

donde en estas fórmulas hemos denotado
$$L_f(x_n) = \frac{f(x_n) f''(x_n)}{f'(x_n)^2}. \tag{2.10}$$

Por otro lado, también existen métodos que sacrifican una velocidad de convergencia elevada por un menor coste operacional, como por ejemplo el denominado *método de Newton simplificado*:
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)},$$
o, en general, métodos del tipo
$$x_{n+1} = x_n - a_n f(x_n),$$
donde a_n son parámetros que, de alguna forma u otra, aproximan a $1/f'(x_n)$. Este segundo tipo de métodos adquieren especial importancia cuando se considera la formulación más general del método de Newton-Kantorovich (2.4):
$$x_{n+1} = x_n - A_n F(x_n),$$
donde, en este caso, A_n son operadores lineales que aproximan al inverso del operador $F'(x_n)$.

- **Métodos multipunto y métodos de tipo secante.** El método de Newton es un ejemplo de lo que se conoce como método de un paso, en el cual el iterado x_{n+1} se obtiene a partir del anterior, x_n. Existen otros métodos que utilizan la información de un mayor número de pasos anteriores $(x_n, x_{n-1} \dots x_{n-p})$ para definir el iterado x_{n+1}. Sin duda, el más famoso de estos métodos es el conocido como *método de la secante*:

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n).$$

  Nótese que se trata de un método de dos pasos en el que se ha evitado el tener que evaluar la derivada primera $f'(x)$. En líneas generales este tipo de métodos busca reducir el coste operacional, aún a costa de perder velocidad de convergencia. En el caso particular del método de la secante se alcanza una convergencia superlineal (de orden $(1+\sqrt{5})/2)$), sin alcanzar la convergencia cuadrática del método de Newton (se define orden de convergencia en la definición 2.1).

- **Métodos para raíces múltiples.** Es bien conocido que las propiedades del método de Newton cambian cuando se aplica para aproximar raíces múltiples. En concreto, la velocidad de convergencia pasa de cuadrática a lineal. Existen diversas variantes del método de Newton que permiten recuperar la convergencia cuadrática. Por ejemplo, cuando se trata de aproximar una raíz de multiplicidad p, el método

$$x_{n+1} = x_n - p\frac{f(x_n)}{f'(x_n)},$$

  converge cuadráticamente a dicha raíz. Pero este método tiene algunos inconvenientes. De entrada, la multiplicidad de la raíz no es conocida. Además, aunque la convergencia es cuadrática hacia la raíz múltiple considerada, pierde la convergencia cuadrática al aproximar al resto de raíces. Otra posibilidad es aplicar el método de Newton a la función $g(x) = f(x)/f'(x)$. Se obtiene así el siguiente método, atribuido a Schröder (1870):

$$x_{n+1} = x_n - \frac{1}{1 - L_f(x_n)} \frac{f(x_n)}{f'(x_n)},$$

  con $L_f(x)$ definido en (2.10), que conserva la convergencia cuadrática incluso para raíces múltiples. Por último, hay que destacar que la situación se complica cuando se pasa de ecuaciones con una variable al caso multidimensional.

- **Estudio dinámico del método de Newton.** El método de Newton puede considerarse como un sistema dinámico discreto. Como tal, el estudio de las órbitas de los diferentes punto de partida puede dar lugar a toda la casuística propia de este tipo de problemas: convergencia a un punto fijo (que es una solución de la ecuación considerada), aparición de ciclos atractores y repulsores, comportamiento caótico, etc. En el caso particular de aplicar el método de Newton a funciones definidas en el campo complejo,

las gráficas de las cuencas de atracción asociadas a las diferentes soluciones tienen una atractiva estructura fractal, como se puede muestra más adelante en el capítulo 4.

- **La versión continua del método de Newton.** Hace referencia al problema de valor inicial

$$\begin{cases} x'(t) & = & -\dfrac{f(x(t))}{f'(x(t))} \\ x(0) & = & x_0. \end{cases}$$

Dicho problema se puede resolver mediante diversos métodos numéricos. Por ejemplo, cuando se considera el método de Euler explícito con un paso de integración de $h = 1$, se obtiene la misma sucesión obtenida por el método de Newton aplicado a la ecuación $f(x) = 0$. Si se considera un paso de integración $h \in (0, 1]$, se obtiene el denominado *método de Newton amortiguado*:

$$x_{n+1} = x_n - h\frac{f(x_n)}{f'(x_n)}.$$

Si en lugar de un paso de integración h constante, se consideran pasos adaptativos h_n se puede lograr una estrategia de convergencia global.

- **Métodos para el cálculo simultáneo de las raíces de una ecuación (método de Weierstrass).** El problema de encontrar las raíces $\alpha_1, \ldots, \alpha_n$ de un polinomio

$$p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n = \prod_{i=1}^{n}(z - \alpha_i) \tag{2.11}$$

cuyos coeficientes pueden ser, indistintamente, números reales o complejos, puede transformarse, mediante las fórmulas de Cardano, en resolver el siguiente sistema de ecuaciones no lineales:

$$f_i(z_1, \ldots, z_n) = (-1)^i \phi_i(z_1, \ldots, z_n) - a_i = 0, \quad i = 1, \ldots, n, \tag{2.12}$$

donde ϕ_i son las funciones simétricas elementales:

$$\phi_i(z_1, \ldots, z_n) = \sum_{1 \le j_1 < \cdots < j_i \le n} z_{j_1} \cdots z_{j_i}.$$

Resolviendo el sistema (2.12) encontramos simultáneamente todas las raíces del polinomio p. Se puede emplear el método de Newton para resolver dicho sistema. Así, a partir de un vector $Z^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)})$ se construye la siguiente sucesión:

$$Z^{(k+1)} = Z^{(k)} - F'(Z^{(k)})^{-1}F(Z^{(k)}).$$

Es bien conocido que la expresión anterior admite una formulación equivalente, que permite calcular las iteraciones de forma explícita, sin necesidad de trabajar con la matriz $F'(Z^{(k)})$. En concreto,

$$Z^{(k+1)} = Z^{(k)} - \left(p(z_1^{(k)})/\Pi_1^{(k)}, \ldots, p(z_n^{(k)})/\Pi_n^{(k)}\right)^T, \quad i = 1, \ldots, n, \quad k = 0, 1, 2, \ldots$$

donde
$$\Pi_i^{(k)} = \prod_{\substack{j=1 \\ j\neq i}}^{n} (z_i^{(k)} - z_j^{(k)}).$$

Esta expresión también puede escribirse componente a componente:

$$z_i^{(k+1)} = z_i^{(k)} - \frac{p(z_i^{(k)})}{\displaystyle\prod_{\substack{j=1 \\ j\neq i}}^{n} (z_i^{(k)} - z_j^{(k)})}, \quad i = 1, \dots, n, \quad k = 0, 1, 2, \dots$$

Este proceso, inicialmente atribuido a Weierstrass (1.903), recibe su nombre y también el de otros matemáticos que han trabajado sobre él: Durand (1.960), Kerner (1.966), Dochev (1.962), etc. Los nombres de método de Weierstrass o método de Durand-Kerner son bastante habituales en la literatura matemática.

- **Métodos para resolver sistemas de ecuaciones subdeterminados (inversos generalizados).** En esta ocasión se trata de aplicar alguna variante del método de Newton para resolver un sistema del tipo $F(x) = 0$, con $x \in \mathbb{R}^n$ y $F = (f_1, \dots, f_m)$, es decir $F : \mathbb{R}^n \to \mathbb{R}^m$, con $m < n$. Se trata, por tanto, de un sistema con más incógnitas que ecuaciones. En condiciones adecuadas, el conjunto de ceros de F, $V = F^{-1}(0)$, es una subvariedad de $\mathbb{R}^n$ de dimensión $n - m$ que se puede aproximar numéricamente iterando la función
$$N_F(x) = x - F'(x)^\dagger F(x),$$
donde $F'(x)^\dagger$ es el inverso generalizado (o inverso de Moore-Penrose) del operador lineal $F'(x)$. Dada una matriz $A \in \mathbb{R}^{m\times n}$, su inverso generalizado $A^\dagger \in \mathbb{R}^{n\times m}$ es la única matriz que cumple las propiedades $AA^\dagger A = A$, $A^\dagger AA^\dagger A = A^\dagger$, $(AA^\dagger)^T = AA^\dagger$ y $(A^\dagger A)^T = A^\dagger A$, donde A^T denota la matriz traspuesta de A.

- **Métodos para resolver sistemas de ecuaciones sobredeterminados (método de Newton-Gauss).** El problema considerado es el de resolver un sistema de ecuaciones con más ecuaciones que incógnitas. Es ésta una situación que aparece con frecuencia en la práctica, como por ejemplo en el ajuste de una nube de puntos (no alineados) por una recta de regresión. Al no existir una solución del problema, se introduce un nuevo concepto de solución: la solución por mínimos cuadrados. En concreto, si escribimos el sistema como $F(x) = 0$, con $x \in \mathbb{R}^n$ y $F = (f_1, \dots, f_m)$, $(m > n)$, el problema de resolver la ecuación $F(x) = 0$ se transforma en el de minimizar el funcional
$$G(x) = \frac{1}{2}\|F(x)\|^2 = \sum_{j=1}^{m} |f_j(x)|^2.$$
El problema de encontrar el mínimo absoluto de G es, a priori, difícil de resolver. En consecuencia, se suele abordar el problema de encontrar los mínimos locales de G.

Para ello hay que encontrar los puntos críticos de G como soluciones del sistema de ecuaciones $G'(x) = 0$. Cuando se trata de un sistema con el mismo número de ecuaciones que de incógnitas, se puede emplear el método de Newton o cualquiera de los métodos habituales para resolver este tipo de sistemas.

Otra estrategia, introducida por Gauss en 1809, consiste en linealizar el sistema inicial $F(x) = 0$ de la forma $F(x) + F'(x)(y - x) = 0$. Para el caso en el que el operador $F'(x)$ sea inyectivo, la solución por mínimos cuadrados de este sistema admite una expresión de la forma

$$y = x - (F'(x)^T F'(x))^{-1} F'(x)^T F(x).$$

La iteración de este proceso da lugar a lo que se conoce como *método de Newton-Gauss*. Este problema tiene unas características totalmente diferentes a las del caso subdeterminado. En concreto la convergencia del método de Newton-Gauss a un punto límite no garantiza que éste sea una solución del sistema $F(x) = 0$, pudiendo ser también un punto crítico del funcional G definido anteriormente.

- **Optimización con y sin restricciones.** Muy vinculado al punto anterior está el problema de minimizar un funcional del tipo $\min f(x)$, donde f está definido en un subconjunto de $\mathbb{R}^n$. Este problema puede transformarse, en condiciones adecuadas, en resolver la ecuación $\nabla f(x) = 0$. La aplicación del método de Newton a este problema da lugar a la sucesión

$$x_{n+1} = x_n - H_f(x_n)^{-1} \nabla f(x_n),$$

donde ∇f y H_f denotan, respectivamente, al vector gradiente y a la matriz hessiana del funcional f. Además algunos problemas de optimización con restricciones, como los obtenidos cuando las restricciones vienen dadas por un conjunto de igualdades también pueden resolverse por el método de Newton. Por ejemplo, las soluciones del problema

$$\min f(x), \text{ con } g(x) = 0,$$

siendo $f : \mathbb{R}^n \to \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}^m$, se encuentran entre los puntos críticos de la correspondiente función de Lagrange $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ definida por

$$L(x, y) = f(x) + y_1 g_1(x) + \cdots + y_m g_m(x), \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m,$$

donde $y \in \mathbb{R}^m$ es un vector formado por los multiplicadores de Lagrange y $g_1, \ldots, g_m$ son las funciones componentes de g. En este caso, se tendría que aplicar la versión general del método de Newton (2.4) para resolver el sistema de ecuaciones no lineales dado por

$$L'_x(x, y) = 0, \quad L'_y(x, y) = 0.$$

Este tipo de problemas se puede extender a resolver problemas de cálculo de variaciones, minimizando funcionales definidos en un espacio de Hilbert.

- **Métodos para resolver ecuaciones no diferenciables.** En el estudio teórico de la convergencia del método de Newton para resolver una ecuación no lineal $F(x) = 0$ es habitual exigir la diferenciabilidad del operador involucrado F, junto con otro tipo de condiciones. En ocasiones, hay que trabajar con ecuaciones del tipo $F(x) = G(x) + H(x) = 0$ donde G es la parte diferenciable y H la parte no diferenciable del operador inicial F. Para este tipo de problemas se puede usar una variante del método de Newton que sólo toma en consideración la parte diferenciable a la hora de calcular el correspondiente inverso, lo que da lugar al siguiente tipo de iteraciones, formuladas en su versión más general:

$$x_{n+1} = x_n - G'(x_n)^{-1}F(x_n) = x_n - G'(x_n)^{-1}[G(x_n) + H(x_n)].$$

- **Implementaciones computacionales del método de Newton.** Una cosa es el estudio de un método numérico desde el punto de vista teórico y otra muy distinta es su aplicación en determinados problemas reales. Al tratar un problema computacionalmente hay que tener en cuenta factores como el acceso libre a los códigos, el tiempo de cálculo, el almacenamiento de los datos y también el tiempo que necesita el usuario para plantear y resolver el problema. Muchos de los programas de cálculo simbólico o numérico más empleados (Mathematica, Matlab, Maple, Derive, Octave, Maxima, Sage, etc.) tienen incorporadas funciones que calculan de forma aproximada las soluciones de una ecuación no lineal. Dichas funciones usan el método de Newton y sus variantes. Además, se puede acceder a los códigos de algoritmos que implementan el método de Newton en diversos lenguajes de programación (C, C++, FORTRAN, etc.). Actualmente, Internet ofrece un gran número de posibilidades para acceder a este tipo de recursos. Como una pequeña muestra, citamos algunas páginas web desde donde se pueden descargar este tipo de programas: [96], [104], [108], [116], [130], [136], [156].

2.4. Convergencia del método de Newton

A finales del siglo XVII y comienzos del XVIII científicos de la talla de Newton, Raphson y Simpson establecieron las bases de lo que hoy conocemos como método de Newton como herramienta para resolver ecuaciones no lineales concretas. El propio Newton explicaba el procedimiento para encontrar una solución de la ecuación $x^3 - 2x - 5 = 0$ próxima al punto $x_0 = 2$. Unos años más tarde, otros científicos como Mourraille, Fourier o Cauchy [31] se plantearon el problema de buscar las condiciones para que las iteraciones del método de Newton (2.2) converjan a una solución de la ecuación (2.1).

En el estudio de la convergencia del método de Newton y, en general, de cualquier proceso iterativo se distinguen tres grandes teorías:

1. Convergencia local: se imponen condiciones sobre la función $f(x)$ que define la ecuación (2.1) y sobre la solución α de la misma.

2. Convergencia semilocal: se imponen condiciones sobre la función $f(x)$ que define la ecuación (2.1) y sobre el punto de partida x_0 de la sucesión (2.2).

3. Convergencia global: se exigen condiciones sobre la función $f(x)$ que define la ecuación (2.1) sobre un rango de valores donde está definida.

En esta sección presentamos resultados de convergencia para el método de Newton que se enmarcan dentro de cada una de las líneas anteriores. Además, analizaremos algunos otros conceptos relacionados con la idea de convergencia (véase [68], [80], [146] para más información).

Definimos ahora uno de los conceptos más importantes en el estudio de los procesos iterativos, como es el de orden de convergencia.

Definición 2.1. *Sea $\{x_n\}$ una sucesión convergente a un límite x^*. Si existe una constante $C \in (0,1)$ tal que*

$$\lim_{n\to\infty} \frac{|x^* - x_{n+1}|}{|x^* - x_n|} = C,$$

se dice que el orden de convergencia es lineal.

Si, además, existe un número real p y una constante no nula C tal que

$$\lim_{n\to\infty} \frac{|x^* - x_{n+1}|}{|x^* - x_n|^p} = C, \tag{2.13}$$

diremos que $\{x_n\}$ es convergente de orden *p y llamaremos a C la* constante de error asintótico.

Cuando $p = 2$, se dice que la convergencia es cuadrática, y cuando $p = 3$, cúbica. Cuando $1 < p < 2$ el orden de convergencia se dice superlineal.

A menudo, para resolver una ecuación $f(x) = 0$, se utiliza un proceso iterativo $x_{n+1} = \phi(x_n)$, donde ϕ es un funcional que depende de la función f. Si queremos asociar el concepto de orden con el proceso iterativo que genera la sucesión, podemos escribir (2.13) como sigue

$$\lim_{x\to x^*} \frac{|x^* - \phi(x)|}{|x^* - x|^p} = C.$$

Para Schröder [135] el proceso anterior es de orden p si

$$\phi(x^*) = x^*; \quad \phi^{(j)}(x^*) = 0, \ j = 1, 2, \ldots, p-1; \quad \phi^{(p)}(x^*) \neq 0.$$

Esta definición solamente es válida para procesos iterativos definidos por funciones de una variable con p derivadas continuas. Una generalización de esta definición para sistemas de ecuaciones puede verse en [146].

Cuando un método iterativo tiene orden de convergencia p, se tiene que en cada iteración se multiplica aproximadamente por p, el número de cifras exactas con las que se aproxima la solución. De esta forma, un algoritmo con orden de convergencia alto converge más rápidamente que un algoritmo de orden menor.

Un estudio más detallado del orden de convergencia puede verse en [109], donde se pone de manifiesto la relación entre el orden y la velocidad de convergencia de la sucesión, así como la independencia de la norma elegida en la definición de orden para el caso de operadores definidos en los espacios $\mathbb{R}^n$. En concreto, consideramos dos procesos iterativos ϕ_1 y ϕ_2 con el mismo límite x^* y órdenes p_1 y p_2, respectivamente. Si $p_1 < p_2$ la sucesión generada por ϕ_2 converge a x^* más rápidamente que la generada por ϕ_1. Además esto ocurre con cualquier norma. Si $p_1 = p_2$, entonces el proceso con menor constante de error asintótico es el más rápido, aunque en este caso no se puede hablar de independencia de la norma elegida.

En un proceso iterativo $x_{n+1} = \phi(x_n)$, el funcional ϕ depende de la función f. Por tanto, el orden del proceso puede variar para distintas clases de funciones f. Así, para hablar con precisión deberíamos decir que un proceso es de al menos un cierto orden. Esta ambigüedad queda suprimida en otras definiciones de orden, como las que dan Ortega y Rheinboldt [109].

La multiplicidad de una raíz puede alterar el orden de un proceso iterativo. Así, cuando decimos que un proceso iterativo es de un cierto orden p, se entiende que esto es cierto para los ceros de una misma multiplicidad. Mientras no se diga lo contrario, supondremos que todos los ceros de las funciones que tratamos son de multiplicidad uno.

2.4.1. Convergencia local del método de Newton

El teorema que presentamos a continuación garantiza la convergencia del método de Newton para puntos de partida suficientemente próximos a la solución. Además, en el caso de raíces simples, se demuestra también que la convergencia es cuadrática. En el mismo se introducen dos elementos, la función de iteración del método de Newton, $N_f(x)$, y su derivada, $L_f(x)$, las cuales toman la forma siguiente:

$$N_f(x) = x - \frac{f(x)}{f'(x)}, \tag{2.14}$$

$$L_f(x) = \frac{f(x)f''(x)}{f'(x)^2}. \tag{2.15}$$

Teorema 2.1. *Sea f una función dos veces diferenciable en un entorno I de una solución α de $f(x) = 0$. Supongamos que $f'(x) \neq 0$ para $x \in I$. Entonces existe un valor $r > 0$ tal que si $x_0 \in (\alpha - r, \alpha + r)$ la sucesión definida por el método de Newton (2.2),*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

converge a α. Además, la convergencia es cuadrática, con constante de error asintótico

$$C = \frac{f''(\alpha)}{2f'(\alpha)}.$$

Demostración. Teniendo en cuenta que $N_f'(t) = L_f(t)$, podemos escribir

$$\alpha - x_{n+1} = \alpha - N_f(x_n) = N_f(\alpha) - N_f(x_n) = \int_{x_n}^{\alpha} L_f(t)\, dt.$$

Como $L_f(\alpha) = 0$ y $L_f(t)$ es una función continua en I, existe $r \in \mathbb{R}$ tal que si $t \in (\alpha - r, \alpha + r)$, entonces $|L_f(t)| \leq \gamma < 1$.

De esta forma, si $x_0 \in (\alpha - r, \alpha + r)$, entonces

$$|\alpha - x_1| \leq \left| \int_{x_0}^{\alpha} L_f(t)\, dt \right| \leq \gamma |\alpha - x_0| < r.$$

Por tanto, $x_1 \in (\alpha - r, \alpha + r)$.

Repitiendo el razonamiento de forma inductiva, si $x_n \in (\alpha - r, \alpha + r)$, entonces

$$|\alpha - x_{n+1}| \leq \left| \int_{x_n}^{\alpha} L_f(t)\, dt \right| \leq \gamma |\alpha - x_n| < r,$$

y, por tanto, $x_{n+1} \in (\alpha - r, \alpha + r)$.

Además,

$$|\alpha - x_{n+1}| \leq \gamma |\alpha - x_n| \leq \gamma^2 |\alpha - x_{n-1}| \leq \cdots \leq \gamma^{n+1} |\alpha - x_0|.$$

Luego, para todo $x_0 \in (\alpha - r, \alpha + r)$,

$$\lim_{n \to \infty} x_{n+1} = \alpha.$$

La convergencia cuadrática del método de Newton se deduce del hecho de que, para raíces simples $N_f(\alpha) = \alpha$, $N_f'(\alpha) = 0$ y $N_f''(\alpha) = f''(\alpha)/f'(\alpha)$. Notemos que, en este caso, la constante de error asintótico para el método de Newton es

$$C = \frac{f''(\alpha)}{2f'(\alpha)}.$$

□

El teorema anterior nos garantiza la existencia de un radio r para el cual se asegura la convergencia del método de Newton si empezamos a iterar en puntos x_0 cuya distancia a la solución sea menor que dicho r. Esta misma idea se puede generalizar al método de Newton-Kantorovich (2.4) e, incluso, se puede dar una estimación de dicho radio de accesibilidad. Para ello, consideramos ahora una ecuación

$$F(x) = 0, \quad F : D \subseteq X \to Y, \tag{2.16}$$

siendo F un operador diferenciable en el sentido de Fréchet, que está definido en un dominio D de un espacio de Banach X y con valores en otro espacio de Banach Y. Denotamos x^* a una raíz simple de F, es decir, tal que $F'(x^*)$ tiene inverso. Supongamos que x^* está contenida en D. Un resultado clásico que se puede encontrar en [124] asegura que si F, además de estas condiciones, cumple la condición de Lipschitz:

$$\|F'(x^*)^{-1}[F'(x) - F'(y)]\| \leq K\|x - y\|, \quad x, y \in D, \tag{2.17}$$

entonces el método de Newton-Kantorovich (2.4) converge a x^* para todo $x_0 \in B(x^*, r) = \{x \in D : \|x - x_0\| \leq r\}$, donde

$$r = \frac{2}{3K}.$$

Dependiendo de las condiciones exigidas al operador, se pueden obtener nuevas estimaciones para el radio de accesibilidad r. Por ejemplo, en [6] se aumenta dicho radio sin más que precisar un poco más la condición (2.17). De hecho, se obtiene el siguiente resultado.

Teorema 2.2. *Sea $F : D \subseteq X \to Y$ un operador diferenciable Fréchet definido en un dominio D que contiene a una raíz simple x^* de F. Supongamos que, además de la condición (2.17), se sabe también que*

$$\|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq L\|x - x^*\|, \quad x \in D, \tag{2.18}$$

Entonces, el método de Newton-Kantorovich $\{x_n\}$ $(n \geq 0)$ generado por (2.4) converge de forma cuadrática a x^* para todo punto de partida x_0 en el interior de la bola $B(x^*, r_1) = \{x \in D : \|x - x_0\| \leq r_1\}$, donde

$$r_1 = \frac{2}{K + 2L}.$$

Demostración. La demostración se basa en el hecho de que la condición (2.18), junto con el lema de Banach sobre inversión de operadores [77], garantizan que el operador lineal $F'(x)$ tiene inverso para todo $x \in B(x^*, r_1)$. En efecto, como

$$\|F'(x^*)^{-1}[F'(x) - F'(x^*)]\| \leq Lr_1 = \frac{2L}{K + 2L} < 1.$$

se tiene que

$$\|F'(x)^{-1}F'(x^*)\| \leq \frac{1}{1 - 2L/(K + 2L)} = \frac{K + 2L}{K}, \quad x \in B(x^*, r_1).$$

Por hipótesis, x_0 está en el interior de la bola $B(x^*, r_1)$. Supongamos que también lo están los puntos x_k generados por el método de Newton, para $k = 0, 1, 2, \ldots, n$. Veamos que x_{n+1} también lo está. En efecto,

$$\begin{aligned} x_{n+1} - x^* &= F'(x_n)^{-1}\left[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)\right] \\ &= \int_0^1 F'(x_n)^{-1}[F'(x_n + t(x^* - x_n)) - F'(x_n)](x^* - x_n)\, dt \end{aligned}$$

y, por tanto,

$$\begin{aligned}\|x_{n+1} - x^*\| &\leq \|F'(x_n)^{-1}F(x^*)\|\|x_n - x^*\|^2 \int_0^1 Kt\,dt \\ &\leq \frac{K+2L}{2}\|x_n - x^*\|^2 = \frac{1}{r_1}\|x_n - x^*\|^2.\end{aligned}$$

Por lo tanto,

$$\|x_n - x^*\| \leq \frac{1}{r_1}\|x_{n-1} - x^*\|^2 \leq \frac{1}{r_1}\frac{1}{r_1^2}\|x_{n-2} - x^*\|^{2^2} \leq \cdots \leq r_1\left(\frac{\|x_0 - x^*\|}{r_1}\right)^{2^n},$$

lo que garantiza la convergencia cuadrática de la sucesión x_n a x^*. □

2.4.2. Convergencia semilocal del método de Newton

Cauchy, en 1829, fue el primero en establecer un resultado de convergencia para el método de Newton en el cuál no se asumía la existencia de solución [151]. En su lugar, Cauchy exigía condiciones sobre el punto de partida x_0 de la sucesión (2.2). El resultado de Cauchy está en la base de lo que algo más de un siglo después, el matemático soviético L. V. Kantorovich, utilizara para estudiar el método de Newton (2.4) para operadores definidos entre dos espacios de Banach.

Teorema 2.3. *Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ diferenciable. Supongamos que $f'(x_0) \neq 0$. Denotamos $h_0 = -f(x_0)/f'(x_0)$, $x_1 = x_0 + h_0$, $J_0 = [x_1 - |h_0|, x_1 + |h_0|]$ y $M = \sup_{x\in J_0} |f''(x)|$. Si*

$$2\left|\frac{f(x_0)M}{(f'(x_0))^2}\right| < 1$$

entonces, la ecuación $f(x) = 0$ tiene una única solución en J_0 y el método de Newton (2.2) empezando en x_0 converge a dicha solución.

Nosotros omitimos aquí la demostración de este resultado, ya que éste puede considerarse como la versión escalar del teorema de Kantorovich, que enunciamos unas líneas más abajo. Además, el lector interesado puede encontrar la demostración de este resultado en [112].

En 1948, L.V. Kantorovich probó la convergencia del método para operadores definidos en espacios de Banach. En esta primera demostración, Kantorovich definía un conjunto de sucesiones reales que controlaban el comportamiento de la sucesión definida en espacios de Banach. Un año más tarde, en 1949, el propio Kantorovich presentó una nueva demostración usando por primera vez sucesiones mayorizantes.

Se dice que una sucesión de números reales $\{t_n\}$ mayoriza a una sucesión $\{x_n\}$ definida en un espacio de Banach X si

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \quad n \geq 0.$$

El interés de las sucesiones mayorizantes es que de su convergencia se puede deducir la convergencia de la sucesión en el espacio de Banach. En efecto, si $\{t_n\}$ converge a t^*, entonces existe $x^* \in X$ de manera que la sucesión $\{x_n\}$ converge a x^* y además

$$\|x^* - x_n\| \leq t^* - t_n, \quad n \geq 0.$$

Esta desigualdad nos permite obtener cotas del error para la sucesión definida en espacios de Banach en términos de su correspondiente mayorizante.

Consideramos ahora una ecuación

$$F(x) = 0 \tag{2.19}$$

donde F es un operador diferenciable Fréchet definido entre dos espacios de Banach X e Y. El método de Newton para resolver (2.19) se define como

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \quad n \geq 0. \tag{2.20}$$

Los resultados que vamos a exponer a continuación dan condiciones suficientes (que se conocen como *condiciones de Kantorovich*) para que los términos de la sucesión $\{x_n\}$ definida en (2.20) estén bien definidos (es decir, para cada $n \in \mathbb{N}$ exista el inverso del operador $F'(x_n)$). Además, aseguran la convergencia de $\{x_n\}$ a una solución x^* de la ecuación (2.19), proporcionan acotaciones del error cometido en cada paso y determinan los dominios donde la solución está localizada y donde es única.

Teorema 2.4 (Kantorovich, [77]). *Sea $F : \Omega \subseteq X \to Y$ un operador diferenciable Fréchet en un conjunto convexo y abierto Ω de un espacio de Banach X. Supongamos también que se cumplen las siguientes condiciones:*

(i) *Existe un punto $x_0 \in \Omega$ donde está definido el operador $\Gamma_0 = F'(x_0)^{-1}$;*

(ii) $\|\Gamma_0(F'(x) - F'(y))\| \leq \gamma\|x - y\|$, $x, y \in \Omega$, $\gamma \geq 0$;

(iii) $\|\Gamma_0 F(x_0)\| \leq \beta$;

(iv) $h = \beta\gamma < 1/2$;

(v) $S = \{x : \|x - x_0\| \leq t^*\} \subseteq \Omega$ *donde*

$$t^* = \frac{1 - \sqrt{1 - 2h}}{h}\beta.$$

Entonces:

(1) *La sucesión de Newton (2.20) está bien definida y es convergente a una solución x^* de la ecuación (2.19).*

(2) *La solución está contenida en la bola de centro x_0 y radio t^*, $\overline{B(x_0,t^*)}$, y es única en $B(x_0,t^{**})$ siendo*

$$t^{**} = \frac{1+\sqrt{1-2h}}{h}\beta.$$

(3) *Se tienen las siguientes cotas del error:*

$$\|x^* - x_n\| \leq (t^{**} - t^*)\frac{\theta^{2^n}}{1-\theta^{2^n}}, \quad \theta = \frac{t^*}{t^{**}} < 1.$$

Demostración. El estudio de la convergencia del método de Newton lo haremos en tres etapas:

1. La sucesión está bien definida y es convergente a una solución de $F(x) = 0$.

2. Obtención de cotas del error.

3. Dominio de unicidad de solución.

En primer lugar, veamos que la sucesión está bien definida y es convergente a una solución de la ecuación $F(x) = 0$. Así, observamos que dado $x \in S$,

$$\|I - \Gamma_0 F'(x)\| = \|\Gamma_0[F'(x) - F'(x_0)]\| \leq \gamma\|x - x_0\|$$

$$\leq \gamma t^* = 1 - \sqrt{1-2h} < 1.$$

Entonces, por el lema de Banach sobre inversión de operadores, deducimos que existe $[\Gamma_0 F'(x)]^{-1}$ y además

$$\|[\Gamma_0 F'(x)]^{-1}\| \leq \frac{1}{1-\gamma\|x-x_0\|} \leq \frac{1}{1-\gamma t^*}. \tag{2.21}$$

En consecuencia, tenemos que el operador $F'(x)^{-1}$ existe para todo $x \in S$ y así, la función $G(x) = x - F'(x)^{-1}F(x)$ que genera el método de Newton está bien definida para todo $x \in S$.

En particular, como $x_0 \in S$, $G(x_0) = x_1$ está bien definido y además

$$\|x_1 - x_0\| = \|\Gamma_0 F(x_0)\| \leq \beta \leq t^* = \frac{1-\sqrt{1-2h}}{h}\beta,$$

ya que $h \leq 1 - \sqrt{1-2h}$ para todo $h \in \mathbb{R}$. Por tanto, $x_1 \in S$ y podemos definir $x_2 = G(x_1)$. Además, teniendo en cuenta (2.21),

$$\|x_2 - x_1\| = \|F'(x_1)^{-1}F(x_1)\| \leq \|[\Gamma_0 F(x_1)]^{-1}\|\|\Gamma_0 F(x_1)\|$$

$$\leq \frac{1}{1-\gamma\|x_1-x_0\|}\|\Gamma_0 F(x_1)\|.$$

Ahora bien, de (2.20) se deduce que

$$F'(x_0)(x_1 - x_0) + F(x_0) = 0,$$

luego

$$F(x_1) = F(x_1) - F(x_0) - F'(x_0)(x_1 - x_0) = \int_{x_0}^{x_1} [F'(x) - F'(x_0)]\,dx$$
$$= \int_0^1 [F'(x_0 + t(x_1 - x_0)) - F'(x_0)](x_1 - x_0)\,dt.$$

Así, tomando normas en la igualdad anterior y teniendo en cuenta que $F'(x)$ satisface una condición de Lipschitz, se llega a

$$\|\Gamma_0 F(x_1)\| = \left\| \int_0^1 \Gamma_0[F'(x_0 + t(x_1 - x_0)) - F'(x_0)](x_1 - x_0)\,dt \right\|$$
$$\leq \gamma \|x_1 - x_0\|^2 \int_0^1 t\,dt = \frac{\gamma}{2}\|x_1 - x_0\|^2.$$

En definitiva,

$$\|x_2 - x_1\| \leq \frac{\gamma}{2}\frac{\|x_1 - x_0\|^2}{1 - \gamma\|x_1 - x_0\|} \leq \frac{h\beta}{2(1-h)}.$$

Para encontrar una sucesión que mayorice a $\{x_n\}$, vamos a considerar el polinomio

$$p(t) = \frac{\gamma}{2}t^2 - t + \beta.$$

Es inmediato comprobar que este polinomio cumple las condiciones de Kantorovich. Además, el método de Newton

$$\begin{cases} t_0 &= 0, \\ t_{n+1} &= t_n - \dfrac{p(t_n)}{p'(t_n)} \end{cases}$$

genera una sucesión creciente cuyo límite es t^*, la menor raíz de $p(t) = 0$.

Podemos observar que

$$t_1 - t_0 = -\frac{p(0)}{p'(0)} = \beta,$$

luego $\|x_1 - x_0\| \leq \beta = t_1 - t_0$.

También es cierto que

$$t_2 - t_1 = -\frac{p(t_1)}{p'(t_1)} = \frac{h\beta}{2(1-h)}$$

y, por tanto, $\|x_2 - x_1\| \leq t_2 - t_1$.

¿Será esto cierto en general?, es decir, ¿se cumplirá que

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \tag{2.22}$$

para todo $n \geq 0$?

Supongamos que $x_0, \ldots, x_k \in S$ y que

$$\|x_{i+1} - x_i\| \leq t_{i+1} - t_i, \quad i = 0, \ldots, k-1.$$

Entonces, usando (2.21),

$$\|x_{k+1} - x_k\| = \|F'(x_k)^{-1}F(x_k)\| \leq \|[\Gamma_0 F'(x_k)]^{-1}\|\|\Gamma_0 F(x_k)\|$$

$$\leq \frac{1}{1-\gamma\|x_k - x_0\|}\|\Gamma_0 F(x_k)\|. \tag{2.23}$$

Procediendo como en el caso $k = 1$, llegamos a que

$$F(x_k) = F(x_k) - F(x_{k-1}) - F'(x_{k-1})(x_k - x_{k-1}) = \int_{x_{k-1}}^{x_k} [F'(x) - F'(x_k)]\, dx$$

$$= \int_0^1 [F'(x_k + t(x_k - x_{k-1})) - F'(x_k)](x_k - x_{k-1})\, dt.$$

Tomando normas,

$$\|\Gamma_0 F(x_k)\| \leq \frac{\gamma}{2}\|x_k - x_{k-1}\|^2 \leq \frac{\gamma}{2}(t_k - t_{k-1})^2. \tag{2.24}$$

Luego, por (2.23), y teniendo en cuenta

$$\|x_k - x_0\| \leq \|x_k - x_{k-1}\| + \cdots + \|x_1 - x_0\| \leq t_k - t_{k-1} + \cdots + t_1 - t_0 = t_k,$$

se deduce que

$$\|x_{k+1} - x_k\| \leq \frac{\gamma}{2}\frac{(t_k - t_{k-1})^2}{1 - \gamma t_k}.$$

Por otra parte,

$$t_{k+1} - t_k = -\frac{p(t_k)}{p'(t_k)} = \frac{p(t_k)}{1 - \gamma t_k}.$$

Además,

$$p(t_k) = p(t_{k-1}) + p'(t_{k-1})(t_k - t_{k-1}) + \frac{1}{2}p''(t_{k-1})(t_k - t_{k-1})^2 = \frac{\gamma}{2}(t_k - t_{k-1})^2.$$

Con esto, hemos probado que $x_{k+1} \in S$ y

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k.$$

Entonces, como $\{t_n\}$ es convergente y, por tanto, de Cauchy, se sigue que $\{x_n\}$ es de Cauchy. Como X es un espacio de Banach, $\{x_n\}$ es convergente. Denotaremos x^* a su límite.

Haciendo tender k a infinito en (2.24) se deduce que $F(x^*) = 0$, luego x^* es solución de (2.19).

Vamos a obtener cotas del error para la sucesión $\{x_k\}$ en función de las cotas del error para la sucesión $\{t_k\}$. Como $\|x_{k+p} - x_k\| \leq t_{k+p} - t_k$, haciendo $p \to \infty$, tenemos:

$$\|x^* - x_k\| \leq t^* - t_k.$$

Las conocidas como fórmulas de Ostrowski [112] nos permiten expresar el error $t^* - t_k$ en función de las raíces del polinomio p, proporcionando unas expresiones muy cómodas de calcular. Así, sean t^* y t^{**} las soluciones de $p(t) = 0$. Denotamos

$$a_n = t^* - t_n, \quad b_n = t^{**} - t_n.$$

Entonces

$$p(t_n) = \frac{\gamma}{2} a_n b_n; \quad p'(t_n) = -\frac{\gamma}{2}(a_n + b_n)$$

y en consecuencia,

$$a_{n+1} = t^* - t_{n+1} = t^* - t_n + \frac{p(t_n)}{p'(t_n)} = a_n - \frac{a_n b_n}{a_n + b_n} = \frac{a_n^2}{a_n + b_n}.$$

De forma análoga se muestra que,

$$b_{n+1} = \frac{b_n^2}{a_n + b_n}.$$

Por lo tanto,

$$\frac{a_{n+1}}{b_{n+1}} = \left(\frac{a_n}{b_n}\right)^2 = \left(\frac{a_{n-1}}{b_{n-1}}\right)^{2^2} = \cdots = \left(\frac{a_0}{b_0}\right)^{2^{n+1}}.$$

Llamando $\theta = t^*/t^{**} = a_0/b_0$, tenemos que $a_n/b_n = \theta^{2^n}$ y, por consiguiente,

$$t^* - t_n = (t^{**} - t_n)\theta^{2^n} = (t^{**} - t^* + t^* - t_n)\theta^{2^n},$$

$$t^* - t_n = (t^{**} - t^*)\frac{\theta^{2^n}}{1 - \theta^{2^n}}.$$

Además, con esto queda probado que el método de Newton tiene orden al menos dos en espacios de Banach.

Por último, vamos a analizar la unicidad de solución. Sabemos que

$$x^* \in S = \{x : \|x - x_0\| \le t^*\}.$$

¿En qué conjunto podemos asegurar que es la única solución?

Supongamos que y^* es otra solución de (2.19) en

$$\hat{S} = \{x : \|x - x_0\| < r\},$$

¿cuál es el mayor valor de r que podemos tomar?

Se tiene que

$$\begin{aligned} 0 = \Gamma_0[F(y^*) - F(x^*)] &= \left[\int_0^1 \Gamma_0 F'(x^* + t(y^* - x^*))\, dt\right](y^* - x^*) \\ &= M(y^* - x^*). \end{aligned}$$

Si el operador M es inversible, deducimos que $x^* = y^*$. Veamos cuándo podemos probar la existencia de inverso para M. Por el lema de Banach, basta con asegurar que

$$\|I - M\| < 1.$$

Así,

$$\|I - M\| = \left\| \int_0^1 \Gamma_0[F'(x^* + t(y^* - x^*)) - F'(x_0)]\, dt \right\|$$

$$\leq \int_0^1 \gamma \|x^* + t(y^* - x^*) - x_0\|\, dt \leq \int_0^1 \gamma\left((1-t)\|x^* - x_0\| + t\|y^* - x_0\|\right)\, dt$$

$$< \frac{\gamma}{2}[t^* + r].$$

Buscamos el mayor valor de r que cumpla

$$\frac{\gamma}{2}[t^* + r] \leq 1,$$

es decir,

$$r = \frac{2}{\gamma} - t^* = t^{**}.$$

En resumen, hemos probado que x^* es la única solución de (2.19) en la bola

$$\hat{S} = \{x : \|x - x_0\| < t^{**}\}.$$

□

2.4.3. Convergencia global del método de Newton

Existen otro tipo de resultados que no exigen, a priori, condiciones sobre la solución ni sobre el punto de partida. Entre éstos destacan (véanse por ejemplo las referencias [64], [112] o [146]) las conocidas como *condiciones de Fourier*, en las cuales se imponen condiciones sobre una función en un intervalo. No obstante, tenemos que hacer notar que de estas condiciones se deduce la existencia de una única solución luego, aunque sea implícitamente, sí que se están exigiendo condiciones sobre la solución.

Teorema 2.5 (Fourier)**.** *Sea $f : [a,b] \to \mathbb{R}$ una función $f \in \mathcal{C}^2[a,b]$ que cumple las siguientes condiciones:*

1. $f(a)f(b) < 0$.
2. $f'(x) \neq 0$, *para todo* $x \in [a,b]$.
3. $f''(x)$ *no cambia de signo en* $[a,b]$.
4. $\max\{|f(a)/f'(a)|, |f(b)/f'(b)|\} \leq b - a$.

Entonces existe una única raíz α de la ecuación $f(x) = 0$ en $[a,b]$ y la sucesión $\{x_n\}$, definida por (2.2) converge hacia α para cualquier valor inicial $x_0 \in [a,b]$.

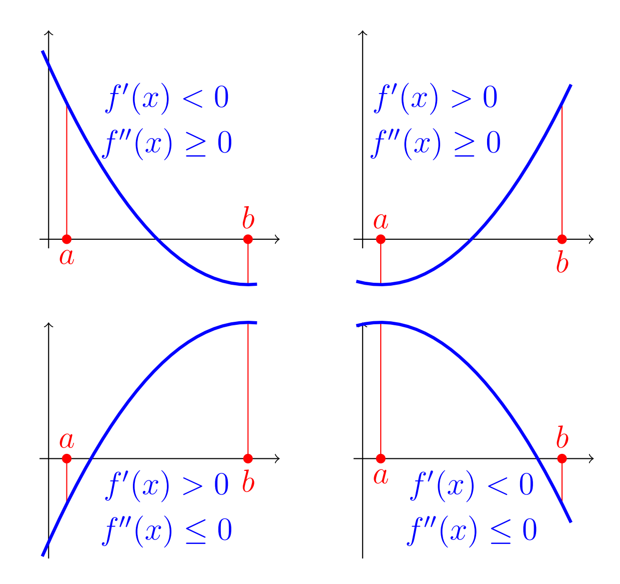


Figura 2.12: Las cuatro representaciones gráficas asociadas a las funciones que cumplen las condiciones del teorema de Fourier.

Demostración. La demostración se basa en probar que el método de Newton (2.2) converge de forma monótona a la solución α. Bajo las condiciones del teorema se pueden presentar cuatro situaciones, que son las que se muestran en la figura 2.12.

Notemos que basta con demostrar la convergencia en una de las cuatro situaciones, pues el resto se puede reducir a un mismo caso mediante cambios de variable o de función, tal y como se hace en [65], por ejemplo. Así, supongamos sin pérdida de generalidad que $f'(x) < 0$ y $f''(x) \geq 0$ en $[a, b]$.

Entonces, si $x_0 \in [a, \alpha]$ se tiene que $\{x_n\}$ es una sucesión creciente y su límite es la raíz α. En concreto, se puede probar por inducción que las desigualdades $x_n < x_{n+1} < \alpha$ son ciertas para $n \geq 0$. Para ello, lo único que hay que tener en cuenta es que en estas condiciones, $f'(x_n) < 0$, $f(x_n) > 0$ y

$$\alpha - x_{n+1} = N_f(\alpha) - N_f(x_n) = N_f'(\gamma)(\alpha - x_n),$$

donde

$$N_f'(\gamma) = \frac{f(\gamma)f''(\gamma)}{f'(\gamma)^2} > 0, \quad \gamma \in (x_0, \alpha).$$

Como $\{x_n\}$ es una sucesión creciente y acotada, tiene un límite que denotamos L. Teniendo en cuenta que

$$L = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \left(x_n - \frac{f(x_n)}{f'(x_n)} \right) = L - \frac{f(L)}{f'(L)},$$

y que $f'(L) \neq 0$, se deduce que $f(L) = 0$ y, por tanto, $L = \alpha$.

Por último si $x_0 \in [\alpha, b]$ se tiene que $x_1 \in [a, \alpha)$ y a partir de aquí la sucesión crece monótonamente hacia la solución. En efecto, por una parte,

$$\alpha - x_1 = -\int_{x_0}^{\alpha} \frac{f'(x) - f'(x_0)}{f'(x_0)}\, dx \geq 0,$$

luego $x_1 \leq \alpha$. Por otra parte, como $f(b) \leq f(x_0)$ y $f'(x_0) \leq f'(b)$, se tiene que

$$f'(b)f(x_0) \leq f'(x_0)f(x_0) \leq f(b)f'(x_0) \Rightarrow \frac{f(x_0)}{f'(x_0)} \leq \frac{f(b)}{f'(b)}.$$

Por tanto, teniendo en cuenta ahora la cuarta de las hipótesis del teorema,

$$b - a \geq \frac{f(b)}{f'(b)} \geq \frac{f(x_0)}{f'(x_0)} = x_0 - x_1.$$

De aquí se deduce que $x_1 \geq a + (x_0 - b) \geq a$, tal y como se quería probar. □

La cuarta condición del teorema anterior garantiza que aunque la primera iteración del método de Newton pueda «saltar» al otro lado de la raíz, en ningún caso se escapa del intervalo $[a, b]$ considerado. Como consecuencia del teorema anterior, se puede dar el siguiente, resultado donde se elimina la cuarta condición y en su lugar aparece una condición sobre el punto de partida que evita que se produzca el «salto» en la primera iteración.

Corolario 2.6. *Sea $f : [a, b] \to \mathbb{R}$ una función $f \in \mathcal{C}^2[a, b]$ que cumple las siguientes condiciones:*

1. $f(a)f(b) < 0$.
2. $f'(x) \neq 0$, *para todo* $x \in [a, b]$.
3. $f''(x)$ *no cambia de signo en* $[a, b]$.

Entonces existe una única raíz α de la ecuación $f(x) = 0$ en $[a, b]$ y la sucesión $\{x_n\}$, definida por (2.2) converge hacia α para cualquier valor inicial $x_0 \in [a, b]$ tal que $f(x_0)f''(x_0) \geq 0$.

2.5. El caso de las raíces múltiples

Es bien conocido [146] que si el punto de partida z_0 está lo suficientemente próximo a una raíz de la ecuación $f(z) = 0$, el método de Newton

$$z_{n+1} = N(z_n) = z_n - \frac{f(z_n)}{f'(z_n)},$$

converge a dicha raíz. Además, se sabe que si la raíz es simple, la sucesión converge cuadráticamente. Sin embargo, si la raíz es múltiple, el orden de convergencia que se obtiene es sólo lineal.

Para el caso de una función f de variable real o compleja, diremos que α es una raíz múltiple o un cero de multiplicidad m de f si

$$f(\alpha) = f'(\alpha) = \cdots = f^{(m-1)}(\alpha) = 0 \quad \text{y} \quad f^{(m)}(\alpha) \neq 0. \tag{2.25}$$

En este caso, se puede escribir

$$f(x) = (x-\alpha)^m g(x), \text{ con } g(\alpha) \neq 0,$$

$$f'(x) = (x-\alpha)^{m-1}(mg(x) + (x-\alpha)g'(x)),$$
$$\frac{f(x)}{f'(x)} = (x-\alpha)\frac{g(x)}{mg(x) + (x-\alpha)g'(x)}.$$

Entonces para valores de x cercanos a α tenemos

$$\frac{f(x)}{f'(x)} = (x-\alpha)\frac{1+\varepsilon(x)}{m}, \quad \lim_{x\to\alpha} \varepsilon(x) = 0.$$

Para $1 \leq q \leq m$ consideramos el proceso

$$x_{n+1} = x_n - q\frac{f(x_n)}{f'(x_n)}. \tag{2.26}$$

Para este método se tiene que

$$x_{n+1} - \alpha = (x_n - \alpha)\left(1 - \frac{q}{m}\right) - \frac{q}{m}(x_n - \alpha)\varepsilon(x_n). \tag{2.27}$$

Por lo tanto, si $q < m$ (situación que ocurre, por ejemplo, en el método de Newton donde $q = 1$),

$$\lim_{n\to\infty} \frac{(x_{n+1} - \alpha)}{(x_n - \alpha)} = 1 - \frac{q}{m} \neq 0.$$

Por consiguiente, se ha perdido la convergencia cuadrática del proceso, pudiéndose garantizar únicamente la convergencia lineal.

Sin embargo, si se toma $q = m$, el valor de la multiplicidad de la raíz y se supone que $f^{(m+1)}(\alpha)$ está definido, entonces el método

$$x_{n+1} = x_n - m\frac{f(x_n)}{f'(x_n)}. \tag{2.28}$$

recupera la convergencia cuadrática. Es más, en este caso,

$$\lim_{n\to\infty} \frac{(x_{n+1} - \alpha)}{(x_n - \alpha)^2} = \frac{f^{(m+1)}(\alpha)}{(m+1)!f^{(m)}(\alpha)}.$$

Llegados a este punto, conviene destacar que si no se puede asegurar la existencia de $f^{(m+1)}(\alpha)$, únicamente se podría garantizar la convergencia superlineal del método correspondiente. En efecto, de la fórmula (2.27) se deduce

$$\lim_{n\to\infty} \frac{(x_{n+1} - \alpha)}{(x_n - \alpha)} = -\lim_{n\to\infty} \frac{q}{m}(x_n - \alpha)\varepsilon(x_n) = 0.$$

Para recuperar la convergencia cuadrática se han desarrollado varias estrategias:

1. Si se conoce que una raíz tiene multiplicidad m, se puede aplicar el método (2.28). Alternativamente, este método se puede obtener aplicando el método de Newton a la función $f(z)^{1/m}$. Se obtiene así una nueva función de iteración:

$$N_m(z) = z - \frac{f(z)^{1/m}}{(1/m)f(z)^{1/m-1}f'(z)} = z - \frac{mf(z)}{f'(z)},$$

que da lugar al conocido como método de Newton relajado. Como hemos indicado unas líneas más arriba, este método converge cuadráticamente (localmente) a la raíz con multiplicidad m. El resto de raíces pasan a ser puntos fijos con multiplicador asociado $1-m$. El mayor problema para la utilización de este método es que se necesita conocer a priori la multiplicidad de la raíz que se pretende aproximar.

2. Otra posibilidad consiste en aplicar el método de Newton a la función racional $f(z)/f'(z)$. Observemos que las raíces de esta función son simples, lo que garantiza la convergencia cuadrática del método de Newton a todas las raíces de $f(z)$. Como contrapartida, observemos que se introducen también los polos que provienen de las raíces de $f'(z)$ que no son raíces de $f(z)$. El nuevo método obtenido, conocido como método de Newton para raíces múltiples o también conocido como método de Schröder, ($\tilde{N}$) tiene la siguiente función de iteración:

$$\tilde{N}(z) = z - \frac{1}{1-L_f(z)}\frac{f(z)}{f'(z)},$$

donde

$$L_f(z) = \frac{f(z)f''(z)}{f'(z)^2}.$$

3. El proceso conocido como aceleración Δ^2 de Aitken es otra forma de acelerar la convergencia del método de Newton para raíces múltiples y, en general, la de cualquier proceso con convergencia lineal. Así, supongamos que $\{z_n\}$ es una sucesión de estas características y que z^* es su límite. Denotamos $e_n = z_n - z^*$. Entonces, por la definición de convergencia lineal, se tiene

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = \lambda \in (0,1).$$

Supongamos que todos los términos de la sucesión de errores $\{e_n\}$ son del mismo signo. Para construir una sucesión que converja a z^* más rápidamente que $\{z_n\}$, tomemos n lo suficientemente grande para poder asegurar que $e_{n+1} \sim \lambda e_n$ y $e_{n+2} \sim \lambda e_{n+1}$. Entonces:

$$z_{n+2} = z^* + e_{n+2} \sim z^* + \lambda e_{n+1} = z^* + \lambda(z_{n+1} - z^*). \tag{2.29}$$

De forma totalmente análoga, también se tiene la equivalencia

$$z_{n+1} \sim z^* + \lambda(z_n - z^*).$$

Restando ambas equivalencias se obtiene una estimación para λ del tipo

$$\lambda \sim \frac{z_{n+2} - z_{n+1}}{z_{n+1} - z_n}. \tag{2.30}$$

Entonces, sustituyendo (2.30) en (2.29) se obtiene una nueva aproximación para z^* del siguiente tipo:

$$z^* \sim \frac{z_{n+2} - \lambda z_{n+1}}{1 - \lambda} \sim z_n - \frac{(z_{n+1} - z_n)^2}{z_{n+2} - 2z_{n+1} + z_n}.$$

El método Δ^2 de Aitken se basa en el hecho de que la sucesión $\{\hat{z}_n\}$ definida por

$$\hat{z}_n = z_n - \frac{(z_{n+1} - z_n)^2}{z_{n+2} - 2z_{n+1} + z_n} \tag{2.31}$$

converge a z^* más rápidamente que $\{z_n\}$. En efecto, en primer lugar, tengamos en cuenta que podemos escribir $e_{n+1} = (\lambda + \delta_n)e_n$, con $\delta_n \to 0$ cuando $n \to \infty$. En consecuencia,

$$e_{n+2} = (\lambda + \delta_{n+1})e_{n+1} = (\lambda + \delta_{n+1})(\lambda + \delta_n)e_n.$$

Ahora bien, como

$$\hat{z}_n - z^* = \frac{e_n e_{n+2} - e_{n+1}^2}{e_{n+2} - 2e_{n+1} + e_n},$$

se deduce que

$$\lim_{n\to\infty} \frac{\hat{z}_n - z^*}{z_n - z^*} = \frac{(\lambda + \delta_{n+1})(\lambda + \delta_n) - (\lambda + \delta_n)^2}{(\lambda + \delta_{n+1})(\lambda + \delta_n) - 2(\lambda + \delta_n) + 1} = 0.$$

Esta última igualdad nos garantiza que la sucesión $\{\hat{z}_n\}$ definida en (2.31) converge a z^* y además, más rápidamente que $\{z_n\}$.

La notación Δ asociada a este proceso tiene su origen en el concepto de *diferencia progresiva* de una sucesión $\{z_n\}$:

$$\Delta z_n = z_{n+1} - z_n, \quad \Delta^k z_n = \Delta^{k-1}(\Delta z_n), \quad k \geq 0.$$

Con esta notación el proceso definido en (2.31) se puede escribir

$$\hat{z}_n = z_n - \frac{(\Delta z_n)^2}{\Delta^2 z_n}.$$

4. Para resolver una ecuación de la forma $f(x) = 0$, donde $f(x)$ es una función de la forma (2.25), el método de Van de Vel consiste en definir dos sucesiones, una que aproxima a la raíz múltiple, α, y la otra a su multiplicidad, m.

   Para ello, se considera el proceso iterativo (2.26) y su correspondiente fórmula del error (2.27). De ésta se deduce que

$$x_{n+1} - \alpha \sim (1 - q/m)(x_n - \alpha)$$

y, por tanto,

$$m \sim q \frac{\alpha - x_n}{x_{n+1} - x_n}.$$

El cociente indicado en la fórmula anterior es, por tanto, una buena aproximación de la multiplicidad m. Sin embargo, tiene el inconveniente de involucrar a la raíz α en su expresión. Una forma de solventar esta dificultad es considerar una aproximación adecuada para α como, por ejemplo, la proporcionada por el método Δ^2 de Aitken (2.31). Así, haciendo

$$\alpha \sim \hat{x}_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}.$$

se tiene que

$$m \sim q \frac{x_n - x_{n+1}}{x_{n+2} - x_{n+1} - (x_{n+1} - x_n)}.$$

Teniendo en cuenta que, a partir de la expresión del proceso iterativo (2.26) se tiene que $x_{n+1} - x_n = -qf(x_n)/f'(x_n)$ y $x_{n+2} - x_{n+1} = -qf(x_{n+1})/f'(x_{n+1})$, podemos escribir

$$m \sim q \frac{-u(x_{n+1})}{u(x_{n+1}) - u(x_n)}, \quad u(x) = \frac{f(x)}{f'(x)}.$$

De esta forma, a partir de unas aproximaciones iniciales x_0 (de la raíz α) y m_0 (de su multiplicidad) tenemos el siguiente proceso iterativo:

$$\left\{ \begin{array}{lcl} m_{n+1} & = & m_n \dfrac{u(x_n)}{u(x_n) - u(x_n - m_n u(x_n))}, \\ x_{n+1} & = & x_n - m_{n+1} u(x_n), \end{array} \right. \qquad u(x) = \frac{f(x)}{f'(x)}, \quad n \geq 0. \tag{2.32}$$

En el artículo [147] se demuestra que la sucesión $\{x_n\}$ definida en el proceso anterior tiene convergencia cuadrática a la raíz múltiple α. Además, la sucesión $\{m_n\}$ converge a la multiplicidad de α como raíz de (2.25). Para una información más completa sobre el algoritmo de Van de Vel, se pueden consultar también las referencias [81] y [148].

2.6. Ejemplos y aplicaciones del método de Newton

2.6.1. Ejemplos históricos

Como ya se puso de manifiesto en la introducción histórica de este capítulo, muchos historiadores consideran al método de Herón para el cálculo de raíces cuadradas como el antecesor del método de Newton. El siguiente ejemplo analiza este problema y su extensión al cálculo de las raíces n-ésimas de un número real positivo.

Ejemplo 2.1. *El conocido como método de Herón es un algoritmo que sirve para aproximar la raíz cuadrada de un número positivo $a > 0$. Aunque fue deducida por el matemático griego*

Herón de Alejandría en el siglo I d. C., también se puede obtener al aplicar el método de Newton para resolver la ecuación

$$f(x) = x^2 - a, \quad a > 0. \tag{2.33}$$

En efecto, en este caso la función de iteración es $N_f(x) = (x^2 + a)/(2x)$ y la sucesión originada por el método de Newton es

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Es sencillo comprobar que dicha sucesión converge a $\sqrt{a}$ para cualquier punto de partida $x_0 > 0$ y que además lo hace con convergencia cuadrática. Ambos hechos permiten obtener buenas aproximaciones de $\sqrt{a}$ con unas pocas iteraciones.

La fórmula de Herón se puede generalizar al cálculo de las raíces k-ésimas de un número real positivo (véase [58]), aplicando en este caso al método de Newton a la función $f(x) = x^k - a$ y obteniendo la sucesión

$$x_{n+1} = \frac{(k-1)x_n^k + a}{kx_n^{k-1}}.$$

Desde el punto de vista formal, la misma fórmula anterior se podría aplicar para el cálculo de las raíces k-ésimas de un número complejo cualquiera.

Tal vez la ecuación no polinómica más conocida, tanto por sus aplicaciones en Astronomía, por su relevancia histórica como por su interés intrínseco es la conocida como ecuación de Kepler. Fue enunciada por el astrónomo alemán Johannes Kepler alrededor del año 1609 en su libro *Astronomia Nova*, donde se establecía que la anomalía x de una órbita elíptica viene dada por la ecuación

$$f(x) = x - e \operatorname{sen} x - M, \tag{2.34}$$

siendo M la anomalía media y e la excentricidad de dicha órbita elíptica. Desde el punto de vista de las interpretaciones astronómicas, tanto x como M son ángulos entre 0 y 2π radianes, mientras que la excentricidad es un valor comprendido entre 0 y 1. Además, debido a la simetría de la elipse respecto a su eje mayor, se puede reducir el estudio a anomalías medias comprendidas entre 0 y π.

Existen muchas maneras de resolver la ecuación de Kepler (el texto de Colwell [34] da buena muestra de ello). En este nos centraremos en su resolución mediante métodos iterativos, tal y como se hace por ejemplo en los artículos de Charles y Tatum [33], Danby [38], [39] o Palacios [113].

Ejemplo 2.2. *Consideremos el método de Newton para aproximar numéricamente una solución de la ecuación de Kepler (2.34). Notemos que dicha ecuación se podría generalizar, desde el punto de vista formal, a una ecuación del tipo*

$$f(x) = x - b - a \operatorname{sen} x,$$

donde a y b son dos parámetros reales o complejos dependiendo de que la variable x sea real o compleja.

La aplicación del método de Newton a la ecuación de Kepler (2.34) da lugar a la siguiente función de iteración:

$$N_f(x) = \frac{M + e(\operatorname{sen} x + x\cos x)}{1 - e\cos x}.$$

Un aspecto clave a la hora de estudiar la convergencia del proceso iterativo $x_{n+1} = N_f(x_n)$ a una solución de la ecuación de Kepler (2.34) es la elección del punto de partida. No se garantiza siempre la convergencia del método de Newton a una solución de la ecuación de Kepler. Por ejemplo, suele ser habitual considerar la anomalía media como punto de partida $x_0 = M$ del método de Newton para resolver la ecuación de Kepler. Pero con esta elección, existen valores de e y de M para los que no se obtiene convergencia, tal y como pone de manifiesto Conway [35] para $e = 0.992$, $M = 0.13\pi$. Sin embargo, Charles y Tatum [33], proponen tomar $E_0 = \pi$, ya que justifican que para ese valor se puede asegurar la convergencia del método de Newton a una solución de la ecuación de Kepler (2.34) para cualquier elección de $M \in [0, \pi]$ y $e \in (0, 1)$. En el cuadro 2.1 se muestra el comportamiento de las sucesiones a las que hacen referencia estos dos ejemplos.

Cuadro 2.1: Importancia de la elección del punto de partida en el comportamiento del método de Newton aplicado a la ecuación de Kepler (2.34), con $e = 0.992$ y $M = 0.13\pi$.

x_0	x_1	x_2	x_3	x_4	x_5	¿Convergencia?
M	4.80602	−1.12977	−0.017804	50.0667	−1760.76	No
π	1.76951	1.44453	1.38508	1.38296	1.38296	Sí

2.6.2. Ejemplos patológicos

Sería una ingenuidad pensar que el método de Newton nos permite aproximar la solución de cualquier ecuación no lineal $f(x) = 0$. Como hemos visto en la sección 2.4, en teoría, se deben comprobar una serie de condiciones para poder garantizar las siguiente cuestiones:

1. Que la sucesión sea convergente.
2. Que el límite sea la raíz.
3. Que la convergencia se produzca en un tiempo «razonable».

Los siguientes ejemplos tienen como objetivo el mostrar que, en ocasiones, la sucesión generada por el método de Newton para resolver una ecuación no lineal $f(x) = 0$, no siempre produce los resultados deseables. Los tres primeros ejemplos han sido tomados de [17].

Ejemplo 2.3. *Analícese el comportamiento del método de Newton aplicado a la función $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = e^{1-x} - 1$.*

Obsérvese que la función anterior tiene una única raíz simple $\alpha = 1$, pues $f'(x) = -e^{1-x}$ y $f'(1) = -1$. Ahora, bien su función de Newton viene dada por

$$N_f(x) = x + 1 - \frac{1}{e^{1-x}}\,. \tag{2.35}$$

Como $\alpha = 1$ es una raíz simple de f, este es un punto fijo superatractor para N_f. La figura 2.13 muestra las gráficas de f y N_f.

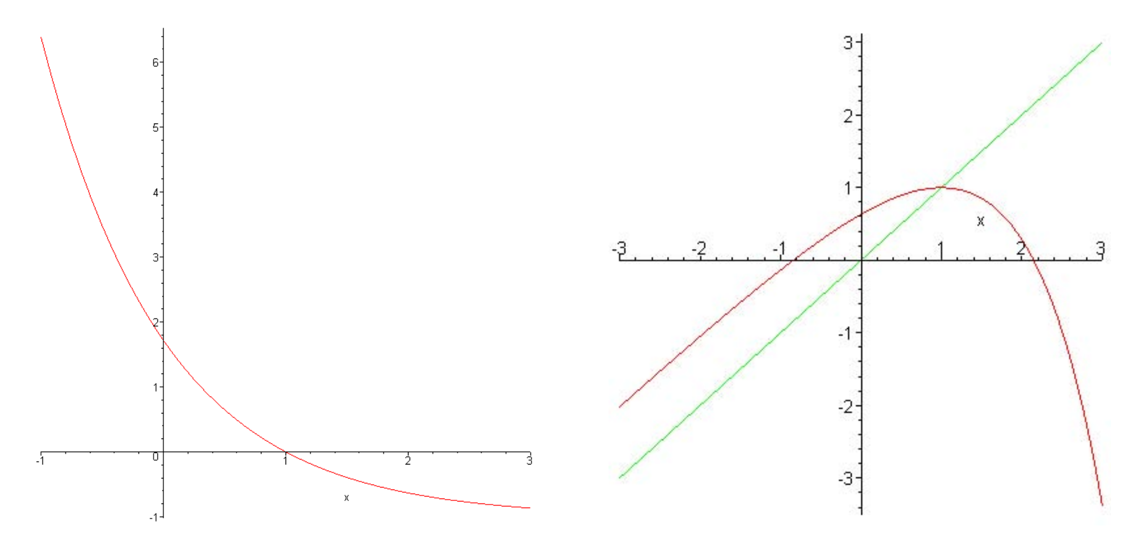


Figura 2.13: Gráficos de $f(x) = e^{1-x} - 1$ y de su correspondiente función de iteración para el método de Newton, $N_f(x)$.

Como $f(x)$ es decreciente y convexa en $\mathbb{R}$, el teorema de Fourier (teorema 2.5) garantiza que si partimos de $x_0 < 1$, la sucesión $x_{n+1} = N_f(x_n)$ es creciente y converge a α. Por otra parte, si partimos de $x_0 > 1$, entonces se tiene que $x_1 = N_f(x_0) < 1$ y a partir de aquí se inicia la convergencia monótona creciente hacia la raíz α. En consecuencia se tiene, en este caso, que la cuenca de atracción de del único cero de f es todo $\mathbb{R}$.

No obstante ocurre el siguiente fenómeno: si partimos de x_0 que es grande en valor absoluto, entonces, x_1 es negativo y de valor absoluto también grande. En consecuencia, se necesitarán muchas iteraciones para aproximar a la raíz α. Para justificar esta afirmación, notemos que $N_f'(x) = 1 - 1/e^{1-x}$, por lo tanto si x es negativo y grande en valor absoluto, se tiene $N_f'(x) \approx 1$, en otras palabras, el

gráfico de N_f es casi paralelo a la diagonal. Entonces, si $x_n < -1$ y es grande en valor absoluto, se tiene que en cada iteración el valor de $x_{n+1} = N_f(x_n)$ es aproximadamente $x_n + 1$, ya que $x_{n+1} = x_n + 1 - 1/(e^{1-x_n})$ y el valor de $1/(e^{1-x_n})$ es muy pequeño. Esto puede apreciarse claramente en la figura 2.14, donde se muestran las iteraciones de la función $N_f(x)$, a partir de $x_0 = 4$.

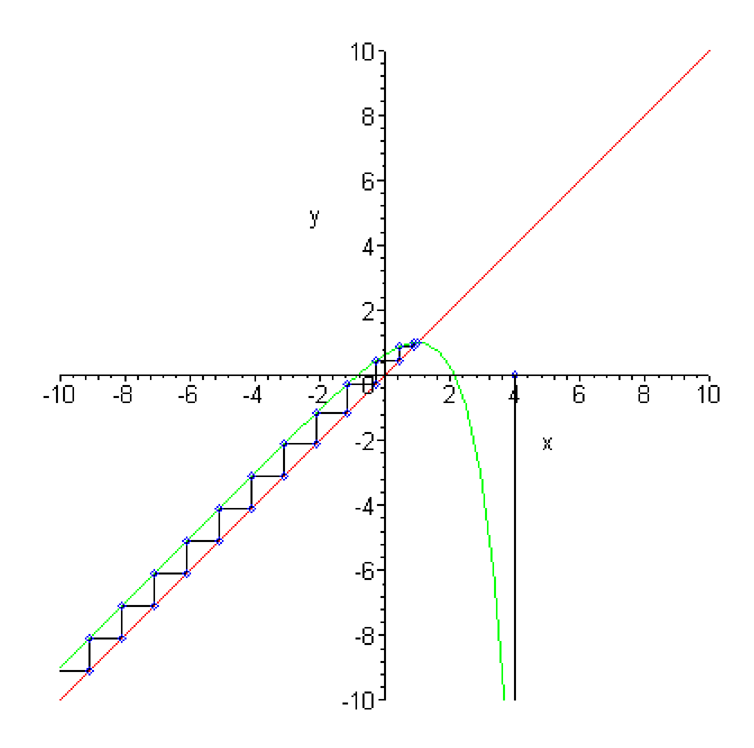


Figura 2.14: Iteraciones de $N_f(x)$, con $x_0 = 4$.

Otra forma de apreciar la convergencia extremadamente lenta del método de Newton en este caso consiste en mostrar algunas iteraciones. En el cuadro 2.2, hemos elegido $x_0 = 10$. Como se ve, todavía estamos muy lejos de α. De hecho, se necesitan muchas más iteraciones para aproximarse a la solución del problema.

Cuadro 2.2: Primeras iteraciones del método de Newton aplicado a la función (2.35) para $x_0 = 10$.

x_0	10	x_5	−8088.0839275753840076
x_1	−8092.0839275753840076	x_6	−8087.0839275753840076
x_2	−8091.0839275753840076	x_7	−8086.0839275753840076
x_3	−8090.0839275753840076	x_8	−8085.0839275753840076
x_4	−8089.0839275753840076	x_9	−8084.0839275753840076

En resumen, aunque la convergencia en este caso es global, es decir, cualesquiera que sea el punto de partida elegido, x_0, la sucesión de los iterados de Newton

$x_{n+1} = N_f(x_n)$ converge a α cuando $n \to \infty$, esta convergencia es extremadamente lenta para valores que no están próximos a la solución. Cabe notar aquí, que lo anterior confirma el hecho de la convergencia cuadrática del método de Newton es sólo local, cosa que evidentemente ocurre en este caso cuando estamos cerca del cero $\alpha = 1$ de nuestra ecuación, como el lector puede verificarlo fácilmente.

Ejemplo 2.4. *Analícese el comportamiento del método de Newton aplicado a la función $f(x) = xe^{-x}$, definida en toda la recta real.*

Es claro que la única solución del problema $f(x) = 0$ es $\alpha = 0$. Tenemos $f'(x) = (1 - x)e^{-x}$ y $c = 1$ es el único punto crítico de f el cual es un máximo. Por lo tanto, N_f tiene una asíntota vertical en c. Estos hechos se ven reflejados en los gráficos de la figura 2.15. Ahora, $f''(x) = (x - 2)e^{-x}$ y $N_f(x) = \dfrac{x^2}{x-1}$. De esto ya podemos obtener las siguientes conclusiones:

- $\lim\limits_{x \to 1^-} N_f(x) = -\infty$.
- $\lim\limits_{x \to 1^+} N_f(x) = +\infty$.
- $\lim\limits_{x \to -\infty} N_f(x) = -\infty$.
- $\lim\limits_{x \to +\infty} N_f(x) = +\infty$.

Como el cero $\alpha = 0$ de f es simple, se tiene que es un punto fijo superatractor. De $N_f'(x) = \dfrac{x(x-2)}{(x-1)^2}$, vemos que $x = 2$ es un punto crítico libre de N_f, es decir, un punto crítico que no es raíz de la ecuación $f(x) = 0$. Dicho punto corresponde a un mínimo en el intervalo $(1, +\infty)$. Notemos que si $x_0 > 1$, entonces la sucesión $x_{n+1} = N_f(x_n)$ es creciente y tiende a $+\infty$ cuando $n \to +\infty$. Por otra parte, la cuenca de atracción de $\alpha = 0$ es el intervalo $(-\infty, 1)$. En este caso, tal y como se precia en los gráficos de la figura 2.16, ocurre un fenómeno similar al del ejemplo anterior, con una convergencia extremadamente lenta cuando se parte de valores negativos que son grandes en valor absoluto.

Experimentalmente, mostramos en el cuadro 2.3 lo dicho arriba, es decir, que si x_0 es negativo y grande en valor absoluto, entonces $x_{n+1} \approx x_n + 1$ y la convergencia se produce muy lentamente.

Ejemplo 2.5. *Analícese el comportamiento del método de Newton aplicado a la función $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = x^{1/3}$.*

En este caso, f no es derivable en $x = 0$ que es la solución de la ecuación $f(x) = 0$. Este hecho nos producirá un efecto no deseado en el método de Newton, ya que para $x \neq 0$, la aplicación de Newton viene dada por $N_f(x) = -2x$. Evidentemente

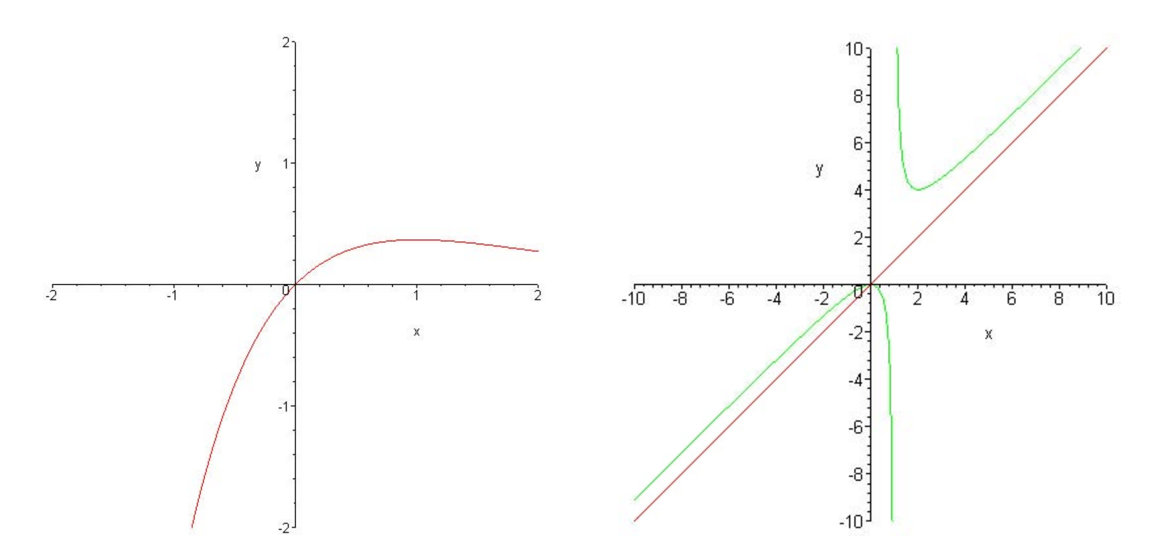


Figura 2.15: Gráficos de $f(x) = xe^{-x}$ y de $N_f(x)$.

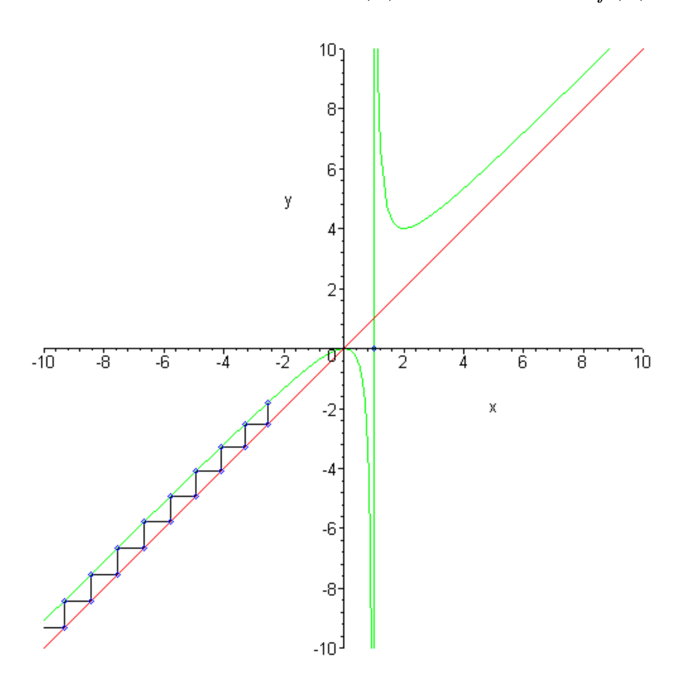


Figura 2.16: Iteraciones de $N_f(x)$, con $x_0 = 0.99$.

esta aplicación tiene a $x = 0$ como punto fijo, pero esta vez es repulsor. De hecho, para todo $x_0 \neq 0$ se tiene que $x_{n+1} = N_f(x_n)$ no converge, de hecho $|N_f(x_n)| \to \infty$ cuando $n \to \infty$.

Ejemplo 2.6. *Analícese el comportamiento del método de Newton para resolver la ecuación* $f(x) = e^{-x} - \operatorname{sen} x = 0$.

Notemos que esta ecuación tiene infinitas soluciones, todas ellas positivas, siendo la menor $\alpha_{-1} \approx 0.588533$. Las restantes son aproximadamente iguales a $\pi, 2\pi, 3\pi, \ldots$. Un cálculo directo muestra que ellas son todas simples. La figu-

Cuadro 2.3: Primeras iteraciones del método de Newton aplicado a la función $f(x) = xe^{-x}$ para $x_0 = -20$ en las que se aprecia una lenta convergencia hacia la solución $\alpha = 0$.

x_0	-20	x_5	-15.263082764788051772
x_1	-19.047619047619047619	x_6	-14.324571721984109704
x_2	-18.097500282773441918	x_7	-13.389826400427451386
x_3	-17.149863156719935957	x_8	-12.459319942056994248
x_4	-16.204959990859412805	x_9	-11.533617900966140257

ra 2.17 muestra la gráfica de $f(x)$ y de su correspondiente función de iteración para el método de Newton, $N_f(x)$.

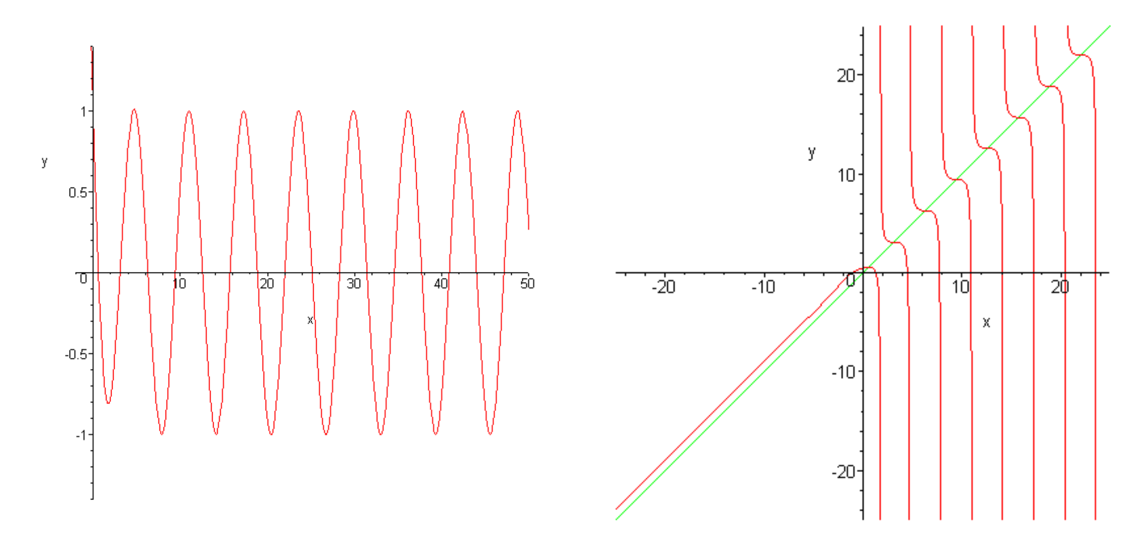


Figura 2.17: Gráfico de $f(x) = e^{-x} - \operatorname{sen}(x)$ y de $N_f(x)$.

El comportamiento dinámico de esta función es complicado, pues tiene infinitos puntos fijos superatractores, infinitos puntos periódicos de período 2, infinitos puntos preperiódicos, etc. Las iteraciones de la función $N_f(x)$ son muy sensibles a la elección de la condición inicial, tal y como se muestra en la figura 2.18, donde se ve que la sucesión $x_{n+1} = N_f(x_n)$, con $x_0 = 11.15$ se aproxima a la raíz que está cerca de 7π. Además, en el cuadro 2.4 se observa que pequeños cambios en el punto de partida provoca grandes cambios en el comportamiento del método de Newton y, en concreto, la convergencia hacia distintos límites.

Ejemplo 2.7. *En este ejemplo, tomado de [67], se muestra el mal comportamiento del método de Newton aplicado a un función que es continua en toda la recta real y que tiene derivada continua en $\{0\}$. En concreto, sea*

$$f(x) = \begin{cases} \pi - 2x\operatorname{sen}(\pi/x) & si \quad x \neq 0, \\ \pi & si \quad x = 0. \end{cases} \tag{2.36}$$

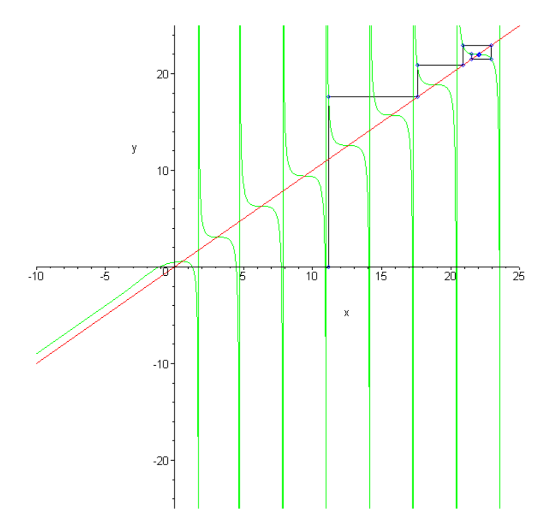


Figura 2.18: Gráfico de algunas iteraciones de $N_f(x)$, con $f(x) = e^{-x} - \operatorname{sen} x$ y $x_0 = 11.15$.

Cuadro 2.4: Iteraciones del método de Newton aplicado a la función $f(x) = e^{-x} - \operatorname{sen} x$, con puntos de partida $x_0 = 11.14$, $x_0 = 11.15$, $x_0 = 11.16$ y $x_0 = 11.17$. Se observa que ligeros cambios en el punto de partida provoca que el límite α de la iteración sea distinto en cada caso.

x_0	11.14	11.15	11.16	11.17
x_1	18.0152	17.5735	17.1864	16.8445
x_2	19.1183	20.8670	6.3865	14.6885
x_3	18.8429	22.9548	6.2847	6.3146
x_4	18.8496	21.5153	6.28505	15.6207
x_5	18.8496	22.0306	6.28505	15.7082
x_6	18.8496	21.9911	6.28505	15.7080
α	$\approx 6\pi$	$\approx 7\pi$	$\approx 2\pi$	$\approx 5\pi$

Entonces la sucesión generada por el método de Newton aplicado a esta función y empezando en $x_0 = 1/2$ converge a 0, pero $f(0) \neq 0$.

Notemos que en este caso, el método de Newton da lugar a la sucesión

$$x_0 = \frac{1}{2}, \quad x_{n+1} = x_n - \frac{\pi x_n - 2x_n^2 \operatorname{sen}(\pi/x_n)}{2\pi \cos(\pi/x_n) - 2x_n \operatorname{sen}(\pi/x_n)}.$$

No es difícil comprobar que $x_n = 1/2^n$ para $n \geq 0$, por lo que, evidentemente, $\lim_{n\to\infty} x_n = 0$ y, sin embargo $f(0) = \pi$, por lo que el límite de la sucesión generada por el método de Newton no es una raíz de la ecuación asociada (véase la figura 2.19).

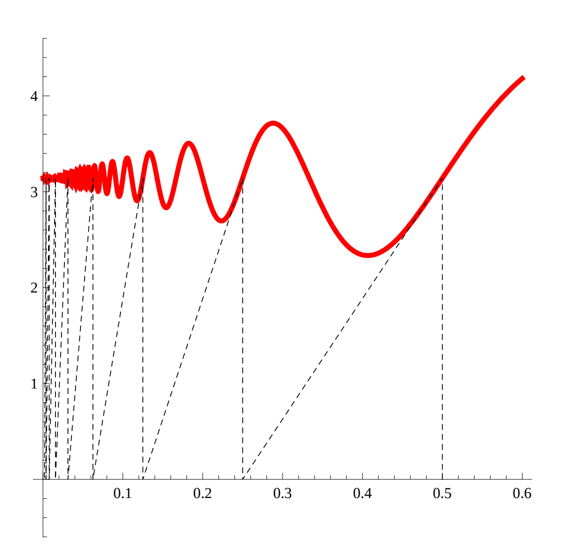


Figura 2.19: El método de Newton aplicado a la función (2.36).

Obsérvese que

$$f'(x_n) = f'(1/2^n) = 2^{n+1}\pi,$$

por lo que $\lim_{n\to\infty} f'(x_n) = \infty$.

Con este ejemplo se muestra que aunque el método de Newton

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

dé lugar a una sucesión convergente a un cierto límite $L \in \mathbb{R}$, no se puede garantizar que $f(L) = 0$. En efecto, si $\lim_{n\to\infty} |f'(x_n)| = \infty$, como ocurre en este ejemplo, entonces puede ocurrir que

$$\lim_{n\to\infty} \frac{f(x_n)}{f'(x_n)} = 0$$

sin que $\lim_{n\to\infty} f(x_n) = 0$.

2.6.3. Sistemas de ecuaciones no lineales

La extensión natural del método de Newton para resolver un sistema de ecuaciones no lineales del tipo $F(x) = 0$, con $x \in \mathbb{R}^n$ y $F = (f_1, \ldots, f_n)$, es decir $F : \mathbb{R}^n \to \mathbb{R}^n$, consiste en construir una sucesión de vectores mediante el siguiente proceso iterativo

$$x^{(k+1)} = x^{(k)} - F'(x^{(k)})^{-1} F(x^{(k)}), \tag{2.37}$$

donde $x^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)}) \in \mathbb{R}^n$ y $F'(x^{(k)})$ es la matriz jacobiana de F formada por las derivadas parciales,

$$F'(x^{(k)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^{(k)}) & \cdots & \frac{\partial f_1}{\partial x_n}(x^{(k)}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^{(k)}) & \cdots & \frac{\partial f_n}{\partial x_n}(x^{(k)}) \end{pmatrix}.$$

En la práctica, en lugar de invertir la matriz jacobiana en cada paso, suele ser más recomendable resolver un sistema de ecuaciones lineales. De esta forma, la expresión alternativa del método de Newton se puede escribir en dos etapas:

$$x^{(k+1)} = x^{(k)} + y^{(k)}, \text{ donde } F'(x^{(k)})y^{(k)} = -F(x^{(k)}).$$

Los siguientes ejemplos tienen como objetivo ilustrar este procedimiento en situaciones sencillas, si entrar a analizar en detalle cuándo se garantiza que la correspondiente sucesión vectorial esté bien definida y sea convergente a una solución del sistema. Para profundizar en estos aspectos se pueden consultar las siguientes referencias: [78], [109], [112], [123].

Ejemplo 2.8. *El sistema de ecuaciones que aparece en este ejemplo ha sido estudiado también en [133]. Usando el método de Newton, resuélvase el sistema de ecuaciones no lineales:*

$$\left\{ \begin{array}{rcl} x - \operatorname{sen} x \cosh y & = & 0 \\ y - \cos x \operatorname{senh} y & = & 0, \end{array} \right. \tag{2.38}$$

partiendo de $(x_0, y_0) = (14, 5)$.

El sistema anterior se puede escribir con notación vectorial de la forma

$$F(x, y) = (f_1(x, y), f_2(x, y))^T$$

donde

$$f_1(x, y) = x - \operatorname{sen} x \cosh y, \quad f_2(x, y) = y - \cos x \operatorname{senh} y.$$

La matriz asociada al operador lineal $F'(x, y)$ (matriz jacobiana) es:

$$F'(x, y) = \begin{pmatrix} 1 - \cos(x)\cosh(y) & -\operatorname{sen}(x)\operatorname{senh}(y) \\ \operatorname{sen}(x)\operatorname{senh}(y) & 1 - \cos(x)\cosh(y) \end{pmatrix}.$$

Por lo tanto, el método de Newton se puede escribir de la forma

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} + \begin{pmatrix} u_k \\ v_k \end{pmatrix},$$

siendo $(u_k, v_k)^T$ la solución del sistema de ecuaciones lineales $F'(x_k, y_k)(u_k, v_k)^T = -F(x_k, y_k)$ o equivalentemente

$$\begin{pmatrix} 1 - \cos(x_k)\cosh(y_k) & -\operatorname{sen}(x_k)\operatorname{senh}(y_k) \\ \operatorname{sen}(x_k)\operatorname{senh}(y_k) & 1 - \cos(x_k)\cosh(y_k) \end{pmatrix} \begin{pmatrix} u_k \\ v_k \end{pmatrix} = -\begin{pmatrix} x_k - \operatorname{sen} x_k \cosh y_k \\ y_k - \cos x_k \operatorname{senh} y_k \end{pmatrix}.$$

En el cuadro 2.5 se muestran los primeros 5 pasos de este procedimiento iterativo para obtener la siguiente solución aproximada del sistema (2.38):

$$(x, y) = (13.9, 3.35221).$$

Cuadro 2.5: Primeros pasos del método de Newton para resolver el sistema (2.38).

k	x_k	y_k
0	14	5
1	13.9697	4.19414
2	13.9318	3.62644
3	13.9059	3.38694
4	13.9001	3.35281
5	13.9000	3.35221

Ejemplo 2.9. *El siguiente ejemplo aparece en [50]. Se trata de encontrar, usando el método de Newton, una solución del sistema de tres ecuaciones no lineales con tres incógnitas siguiente*

$$\left\{\begin{array}{rcl} 3x - \cos(yz) - 0.5 & = & 0, \\ x^2 - 81(y+0.1)^2 + \operatorname{sen}(z) + 1.06 & = & 0, \\ 20z + e^{-xy} + 9.472 & = & 0, \end{array}\right. \tag{2.39}$$

partiendo de $(x_0, y_0, z_0) = (0.1, 0.1, -0.1)$.

La matriz jacobiana es en este caso:

$$\begin{pmatrix} 3 & z\operatorname{sen}(yz) & y\operatorname{sen}(yz) \\ 2x & -162(y+0.1) & \cos(z) \\ -e^{-xy}y & -e^{-xy}x & 20 \end{pmatrix}.$$

La sucesión vectorial obtenida por el método de Newton se muestra en el cuadro 2.6. En la misma se observa la convergencia hacia la solución solución:

$$(x, y, z) = (0.5, 0, -0.5236).$$

Cuadro 2.6: Primeros pasos del método de Newton para resolver el sistema (2.39).

k	x_k	y_k	z_k
0	0.1	0.1	−0.1
1	0.49987	0.0194668	−0.523558
2	0.500014	0.00158853	−0.5236
3	0.5	0.00001238	−0.5236
4	0.5	0	−0.5236

Los sistemas de ecuaciones no lineales pueden verse como un caso particular de una ecuación matricial $F(X) = 0$ siendo X y 0 unas matrices de dimensiones adecuadas. El siguiente ejemplo muestra una de estas ecuaciones, conocida como *ecuación de Riccati algebraica*. Un estudio más detallado de este tipo de ecuaciones y de sus numerosas aplicaciones puede encontrarse en el manual de Lancaster y Rodman [88].

Ejemplo 2.10. *La ecuación de Riccati algebraica es un ecuación del tipo*

$$R(X) = XDX - XA - A^T X - C = 0, \tag{2.40}$$

donde $D, A, C \in \mathbb{R}^{n\times n}$ *son matrices dadas,* A^T *denota la matriz traspuesta de* A *y* $X \in \mathbb{R}^{n\times n}$ *es una matriz incógnita. Supondremos que* D *es una matriz simétrica.*

Para poder aplicar la teoría de Newton-Kantorovich a la ecuación de Riccati, notemos en primer lugar que el operador $R(X)$ resulta ser cuadrático, es decir, $R''(X)$ es un operador bilineal constante. Las expresiones de sus derivadas de Fréchet de primer y segundo orden son

$$R'(X)Y = (XD - A^T)Y + Y(DX - A), \quad R''(X)YZ = YDZ + ZDY.$$

Para una norma matricial $\|\cdot\|$ dada se tiene que

$$\|R''(X)YZ\| \le 2\|D\|\|Y\|\|Z\|,$$

luego $\|R''(X)\| \le 2\|D\|$.

Tomamos como punto de partida una matriz simétrica X_0. Entonces para hallar el inverso de $R'(X_0)$, debemos resolver una ecuación de Liapunov de la forma

$$R'(X_0)Y = (DX_0 - A)^T Y + Y(DX_0 - A) = Z \iff R'(X_0)^{-1}Z = Y.$$

Se sabe (véase [88]) que esta ecuación tiene solución si $DX_0 - A$ es estable, es decir, todos sus valores propios tienen parte real negativa. En este caso, la solución es

$$Y = -\int_0^\infty \exp((DX_0 - A)^T t) Z \exp((DX_0 - A)t)\, dt$$

y la iteración correspondiente al método de Newton queda, en este caso, de la forma

$$X_{n+1} = X_n - \int_0^\infty \exp((DX_n - A)^T t) Z \exp((DX_n - A)t)\, dt.$$

Como consecuencia del teorema de Kantorovich (teorema 2.4) se obtiene un resultado de convergencia para la ecuación de Riccati. Si X_0 es una matriz simétrica tal que $DX_0 - A$ es estable y

$$2\|D\|\|R(X_0)\| \left(\int_0^\infty \|\exp((DX_0 - A)^T t)\|\|\exp((DX_0 - A)t)\|\, dt\right)^2 \le \frac{1}{2},$$

entonces el método de Newton converge a una solución de la ecuación de Riccati algebraica.

Consideramos la ecuación de Riccati (2.40) para las matrices

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

Tomamos como punto de partida la matriz

$$X_0 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Entonces, las cuatro primeras matrices obtenidas aplicando el método de Newton a (2.40) son

$$X_1 = \begin{pmatrix} 1/2 & 1 \\ 1 & 3/2 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 1/4 & 5/6 \\ 5/6 & 2/3 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 1/8 & 53/60 \\ 53/60 & 11/18 \end{pmatrix}, \quad X_4 = \begin{pmatrix} 1/16 & 1013/1080 \\ 1013/1080 & 4549/8100 \end{pmatrix}.$$

Se observa que el método converge a la solución exacta de la correspondiente ecuación que es, en este caso:

$$X^* = \begin{pmatrix} 0 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

2.6.4. Ecuaciones y sistemas con raíces múltiples

Cuando se trata de aproximar una raíz múltiple, el método de Newton tiene un comportamiento diferente al caso general. De entrada, se pierde la velocidad de convergencia cuadrática, pasando únicamente a convergencia lineal. En el caso escalar, la multiplicidad de una raíz α se presenta cuando $f(\alpha) = f'(\alpha) = 0$. En el caso de sistemas de ecuaciones lo que ocurre es que la matriz jacobiana $F'(x)$ no tiene inversa en la solución del sistema. En este apartado mostramos algunos ejemplos que usan variantes del método de Newton para ecuaciones con raíces múltiples. Un análisis en mayor profundidad de este tipo de problemas se puede ver en [86], [81], [122], [147] o [148].

Ejemplo 2.11. *Para ilustrar el comportamiento del método de Van de Vel (2.32) se considera el conjunto de funciones test que aparece en [147]:*

$$f(x) = e^x(x^2 - 1)^m, \quad m = 1, 5, 9, 7/5, \pi.$$

Tomamos como puntos de partida $m_0 = 0.99$ y $x_0 = 1.5$ para obtener los datos que se muestran en el cuadro 2.7. En el ejemplo se aprecia que ambas sucesiones $\{x_n\}$ y $\{m_n\}$ tienen una rápida convergencia hacia la raíz múltiple $\alpha = 1$ y su multiplicidad m. Se ha seguido la estrategia de tomar como multiplicidad inicial un valor próximo a 1, lo cual puede ser razonable si no se tiene ningún conocimiento a priori de la multiplicidad de la raíz. Ahora bien, la convergencia de este tipo de procesos está fuertemente condicionada por tener que partir de un punto x_0 lo suficientemente próximo a la solución buscada. En otro caso, los resultados pueden no ser los esperados.

Cuadro 2.7: El método de Van de Vel (2.32) aplicado a las funciones $f(x) = e^x(x^2 - 1)^m$ con $m = 1, 5, 9, 7/5, \pi$.

$m = 1$		
i	m_i	x_i
1	2.17460	0.86041
2	1.07446	1.05138
3	1.07658	0.99999
4	0.99999	1

$m = 5$		
i	m_i	x_i
1	7.91095	0.89146
2	4.85949	1.00601
3	5.02166	0.99999
4	5	1

$m = 9$		
i	m_i	x_i
1	13.4711	0.90393
2	8.75016	1.00341
3	9.01932	1
4	9	1

$m = 7/5$		
i	m_i	x_i
1	2.79954	0.857903
2	1.43462	1.03504
3	1.46018	0.99997
4	1.39996	1

$m = \pi$		
i	m_i	x_i
1	5.30876	0.87835
2	3.06698	1.01082
3	3.17022	0.99999
4	3.14158	1

Ejemplo 2.12. *En este ejemplo, tomado de [122], se analiza el comportamiento del método de Newton para resolver el sistema de ecuaciones*

$$\left\{ \begin{array}{rcl} x^2 - xy + y^2 + x & = & 2, \\ 3x^2 + 2xy + 2y & = & 7. \end{array} \right. \tag{2.41}$$

Un análisis gráfico nos permite situar las soluciones del sistema. Como se ve en la figura 2.20 el sistema tiene 3 soluciones. En dos de ellas la matriz jacobiana es inversible y, por tanto, el método de Newton tiene convergencia cuadrática. Sin embargo la tercera solución está sobre la curva de puntos donde se anula la matriz jacobiana, marcada con un trazo verde discontinuo en la figura 2.20. Se

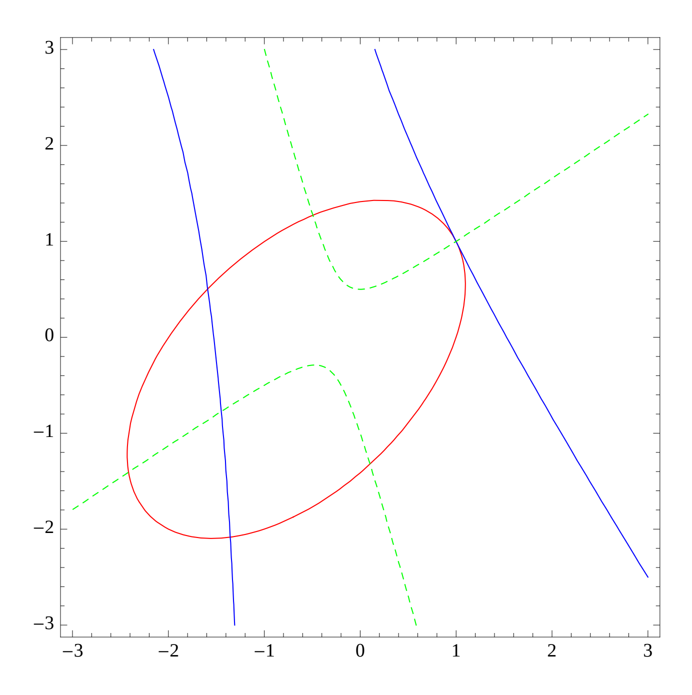


Figura 2.20: Gráfica de las curvas $x^2 - xy + y^2 + x = 2$ (en rojo) y $3x^2 + 2xy + 2y = 7$ (en azul), así como de la curva $10x^2 - 12xy + 6x - 4y^2 - 2y + 2 = 0$ (en verde) formada por los puntos (x, y) donde $\det(F'(x, y)) = 0$.

trata por tanto de una solución singular y el método de Newton tiene un orden de convergencia lineal.

En este caso, la matriz jacobiana es

$$F'(x, y) = \begin{pmatrix} 2x - y + 1 & 2y - x \\ 6x + 2y & 2x + 2 \end{pmatrix}$$

y su determinante $\det(F'(x, y)) = 10x^2 - 12xy + 6x - 4y^2 - 2y + 2$.

Como vemos en el cuadro 2.8, el método de Newton necesita 4 iteraciones para alcanzar cada una de las raíces simples (trabajando con seis dígitos de precisión) mientras que para aproximar la raíz múltiple se necesitan 21 iteraciones.

Cuadro 2.8: El método de Newton para encontrar las tres soluciones del sistema (2.41).

1ª raíz simple			2ª raíz simple			Raíz múltiple		
i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
0	−2	2	0	−2	−3	0	0	0
1	−1.62069	0.982759	1	−1.47222	−2.25000	5	0.96367	1.07267
2	−1.59799	0.566914	2	−1.36138	−2.09177	10	0.99886	1.00227
3	−1.59149	0.506533	3	−1.35601	−2.08406	15	0.99996	1.00007
4	−1.59137	0.505093	4	−1.35600	−2.08404	20	0.99999	1.00000

Ejemplo 2.13. *En [86] se presenta una extensión del método de Newton para raíces múltiples para el caso de sistemas de ecuaciones no lineales. La utilizamos para resolver el sistema de ecuaciones*

$$\left\{ \begin{array}{rcl} x \operatorname{sen} x + y^3 & = & 0, \\ y + x \operatorname{sen} y & = & 0. \end{array} \right. \tag{2.42}$$

El sistema (2.42) tiene en $(x, y) = (0, 0)$ una ecuación singular. En efecto, en este caso, si $F(x, y) = (x \operatorname{sen} x + y^3, y + x \operatorname{sen} y)$ se tiene que

$$F'(x, y) = \begin{pmatrix} x\cos(x) + \operatorname{sen}(x) & 3y^2 \\ \operatorname{sen}(y) & x\cos(y) + 1 \end{pmatrix}, \quad F'(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

y $\det(F'(0, 0)) = 0$.

Si denotamos $f_1(x, y) = x \operatorname{sen} x + y^3$ y $f_2(x, y) = y + x \operatorname{sen} y$, podemos escribir

$$f_1(x, y) = x^2 + \sum_{|\alpha|\geq 3} c_\alpha x^{\alpha_1} y^{\alpha_2}, \quad f_2(x, y) = y + \sum_{|\alpha|\geq 3} d_\alpha x^{\alpha_1} y^{\alpha_2}$$

donde α es un multi-índice, $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$, $|\alpha| = \alpha_1 + \alpha_2$. El menor grado de los términos homogéneos que aparecen en los desarrollos anteriores se denomina *orden* de $(x, y) = (0, 0)$ como cero de f_j, $j = 1, 2$ y se denota por k_j. En nuestro caso, se tiene que $k_1 = 2$, $k_2 = 1$.

Para este problema, (2.42), el método de Newton (2.37) presenta un orden de convergencia lineal para aproximar dicha raíz. Sin embargo, la modificación introducida por Kravanja y Haegemans en [86], dada por

$$X^{(k+1)} = X^{(k)} - F'(X^{(k)})^{-1} \operatorname{diag}(k_1, \ldots, k_n) F(X^{(k)}), \quad X^{(k)} = (x^{(k)}, y^{(k)}), \tag{2.43}$$

recupera el orden de convergencia cuadrático, tal y como se pone de manifiesto en el cuadro 2.9. El método modificado de Kravanja-Haegemans converge en 5 iteraciones, mientras que el método de Newton clásico necesita 18 iteraciones para llegar a la solución con la misma precisión.

2.6.5. Ecuaciones funcionales

El método de Newton, en su versión más general que hemos denominado método de Newton-Kantorovich, permite su aplicación a ecuaciones definidas entre espacios de funciones, como puede ser el caso de ecuaciones diferenciales o ecuaciones integrales. Mostramos a continuación algunos ejemplos.

Cuadro 2.9: El método de Newton (2.37) y la modificación de Kravanja-Haegemans (2.43) para el sistema (2.42).

i	Newton (2.37)	Kravanja-Haegemans (2.43)
0	$(0.75, 0.75)$	$(0.75, 0.75)$
1	$(1.65444, -0.462403)$	$(-0.258806, 0.379649)$
2	$(-0.25613, -0.322061)$	$(-0.146683, -0.048575)$
3	$(0.095719, 0.143380)$	$(0.001126, 0.008402)$
4	$(0.077952, 0.002233)$	$(0.000262, 0.000007)$
4	$(0.038937, 0.000080)$	$(0, 0)$

Ejemplo 2.14. *La ecuación de Chandrasekhar [32] es una ecuación que aparece en diversos problemas físicos (transferencia radiactiva, cinética de gases, etc.). Consideramos el espacio* $X = C[0,1]$ *formado por las funciones continuas en* $[0,1]$ *y dotado de la norma del máximo*

$$\|x\| = \max_{t\in[0,1]} |x(t)|, \quad x \in X.$$

Se trata de encontrar una función $x \in X$ *que cumpla*

$$x(s) = 1 + \lambda x(s) \int_0^1 \frac{s}{s+t} x(t)\, dt, \quad s \in [0,1]. \tag{2.44}$$

Con notación de operadores, el problema puede plantearse como el de resolver la ecuación

$$F(x) = 0$$

siendo F *el operador definido de* X *en* X *tal que*

$$F(x)(s) = x(s) - 1 - \lambda x(s) \int_0^1 \frac{s}{s+t} x(t)\, dt, \quad s \in [0,1].$$

Estudiaremos el caso particular $\lambda = 1/4$.

En primer lugar, como aplicación del teorema de Kantorovich, probaremos la existencia y unicidad de solución de esta ecuación. Posteriormente, encontraremos una aproximación numérica de la misma. Para ello, pasaremos de un problema continuo a un problema discreto utilizando una fórmula de integración numérica (de tipo Gauss-Legendre) para aproximar la integral que aparece en (2.44). Obtenemos así un sistema de ecuaciones (no lineal) que resolvemos por el método de Newton. Las incógnitas son los valores aproximados de la solución en una serie de puntos del intervalo $[0,1]$. Finalmente, a partir de estos valores y mediante un proceso de interpolación se obtiene la aproximación buscada.

Para aplicar el teorema de Kantorovich, necesitamos partir de una función inicial adecuada. Como de (2.44) se deduce que $x(0) = 1$, una elección razonable parece ser $x_0(s) = 1$, para todo $s \in [0, 1]$. Además, hay que calcular el operador derivado F'. Para ello, hay que tener en cuenta que F' actúa de X en $\mathcal{L}(X)$, el espacio de las aplicaciones lineales de X en X. Así, para cada $x \in X$, $F'(x) \in \mathcal{L}(X)$; además $F'(x)y$ es una función continua para cada $x, y \in X$ definida de la siguiente manera:

$$F'(x)y(s) = \lim_{h\to 0} \frac{F(x+hy)(s) - F(x)(s)}{h}$$
$$= y(s) - \frac{1}{4}x(s)\int_0^1 \frac{s}{s+t}y(t)\,dt - \frac{1}{4}y(s)\int_0^1 \frac{s}{s+t}x(t)\,dt, \quad s \in [0,1].$$

Podemos considerar el subconjunto Ω que aparece en el enunciado del teorema de Kantorovich como el propio conjunto X que, por supuesto, es un conjunto convexo. A continuación, calculamos las constantes β y γ.

En primer lugar,

$$\|\Gamma_0 F(x_0)\| \leq \|\Gamma_0\|\|F(x_0)\|;$$

como

$$F(x_0)(s) = -\frac{1}{4}\int_0^1 \frac{s}{s+t}\,dt = -\frac{s}{4}\log\frac{1+s}{s},$$

tenemos que

$$\|F(x_0)\| = \frac{\log 2}{4}.$$

El lema de Banach sobre inversión de operadores nos permite encontrar una cota para $\Gamma_0 = F'(x_0)^{-1}$. Dada una función cualquiera $y \in X$, se tiene que

$$\|[I - F'(x_0)]y\| = \max_{s\in[0,1]} |y(s) - F'(x_0)y(s)|$$
$$= \frac{1}{4}\max_{s\in[0,1]}\left|\int_0^1 \frac{s}{s+t}y(t)\,dt + y(s)\int_0^1 \frac{s}{s+t}\,dt\right| \leq \frac{\log 2}{2}\|y\|,$$

luego

$$\|I - F'(x_0)\| \leq \frac{\log 2}{2} = 0.3465\ldots < 1,$$

con lo que se garantiza que existe $\Gamma_0 = F'(x_0)^{-1}$ y además

$$\|\Gamma_0\| \leq \frac{1}{1 - \|I - F'(x_0)\|} \leq \frac{1}{1 - (\log 2/2)} = 1.5303\ldots$$

En consecuencia,

$$\|\Gamma_0 F(x_0)\| \leq \frac{(\log 2/4)}{1 - (\log 2/2)} = 0.2651\ldots \equiv \beta.$$

De forma muy parecida, podemos deducir que

$$\|\Gamma_0[F'(x) - F'(y)]\| \leq \|\Gamma_0\|\|F'(x) - F'(y)\| \leq \frac{(\log 2/2)}{1 - (\log 2/2)}\|x - y\|.$$

Por tanto,

$$\gamma \equiv \frac{(\log 2/4)}{1-(\log 2/2)} = 2\beta.$$

Así, tenemos que $\alpha = \beta\gamma = 2\beta^2 = 0.1406\ldots < 1/2$, con lo que se cumplen las condiciones del teorema de Kantorovich y se prueba la existencia de solución de (2.44). Aún podemos precisar más, estableciendo las regiones de existencia y unicidad. Para ello, sean

$$t^* = \frac{1-\sqrt{1-2\alpha}}{\gamma} = 0.2870\ldots, \quad t^{**} = \frac{1+\sqrt{1-2\alpha}}{\gamma} = 3.4837\ldots$$

Sabemos que la ecuación de Chandrasekhar tiene una solución x^* localizada en la bola

$$\|x^* - x_0\| \le t^* \qquad (|x^*(s) - 1| \le t^*, \text{ para todo } s \in [0,1]).$$

Además esta solución es única en la bola $\|x^* - x_0\| < t^{**}$.

Pasemos ahora a encontrar la aproximación numérica de la solución de (2.44). Utilizando la fórmula de Gauss-Legendre, podemos hacer

$$\int_0^1 f(t)\,dt \approx \frac{1}{2}\sum_{j=1}^{m} \beta_j f(t_j),$$

donde t_j y β_j son nodos y pesos conocidos (aparecen tabulados para distintos valores de m). En particular, para $m = 8$ son los que aparecen en el cuadro 2.10.

Cuadro 2.10: Pesos (t_j) y nodos (β_j) para la fórmula de Gauss-Legendre para $m = 8$.

j	t_j	β_j	j	t_j	β_j
1	0.01985507	0.10122854	5	0.59171732	0.36268378
2	0.10166676	0.22381034	6	0.762766205	0.31370665
3	0.237233795	0.31370665	7	0.89833324	0.22381034
4	0.40828268	0.36268378	8	0.98014493	0.10122854

Denotando x_i a las aproximaciones de $x(t_i)$, $i = 1,\ldots,8$, llegamos al siguiente sistema de ecuaciones (no lineal):

$$x_i = 1 + \frac{1}{8} x_i t_i \sum_{j=1}^{8} \beta_j \frac{x_j}{t_i + t_j}, \quad i = 1,\ldots,8.$$

Si hacemos $a_{ij} = t_i\beta_j/(8(t_i+t_j))$, podemos escribir el sistema anterior en la forma

$$x_i = 1 + x_i \sum_{j=1}^{8} a_{ij} x_j, \quad i = 1,\ldots,8. \tag{2.45}$$

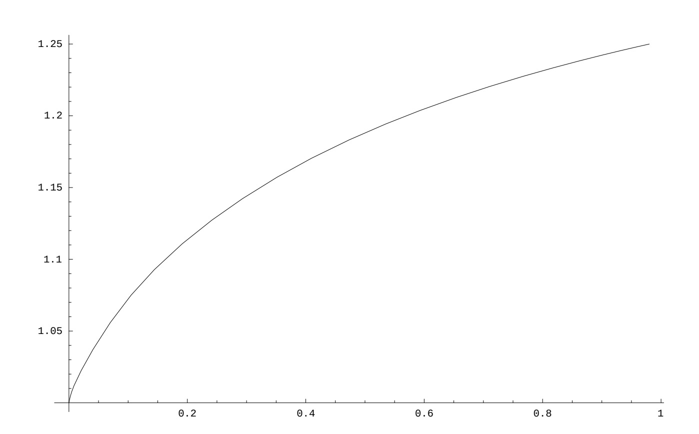


Figura 2.21: Gráfica de la solución aproximada de la ecuación de Chandrasekhar.

En forma matricial, si denotamos $\mathbf{x} = (x_1, \ldots, x_8)^T$ y $\mathbf{1} = (1, \ldots, 1)^T$,

$$\mathbf{x} = \mathbf{1} + \mathbf{x} \odot (A\mathbf{x}),$$

donde $\odot$ representa el producto componente a componente. La solución de este sistema es el siguiente vector:

$$(x_1, \ldots, x_8) = (1.02176, 1.07334, 1.12599, 1.17011, 1.20351, 1.22698, 1.24205, 1.24999).$$

Interpolando la función que pasa por los puntos (t_i, x_i) $i = 1, \ldots, 8$, y sabiendo además, que $x(0) = 1$, obtenemos la gráfica aproximada de la solución, que se muestra en la figura 2.21.

Otra fuente importante de ejemplos donde aplicar el método de Newton son las ecuaciones diferenciales. En este contexto, consideramos un problema de valores en la frontera (PVF) del tipo.

$$\frac{d^2y}{dx^2} + g(y) = 0, \quad y(a) = A, \quad y(b) = B, \tag{2.46}$$

donde $y = y(x)$ es una función real de variable real lo suficientemente derivable en el intervalo $[a, b]$.

Uno de los métodos más empleados para resolver el problema (2.46) es el que se conoce como método de disparo. Sin entrar en muchos detalles (véase [50] o [79] para más información) el método consiste en conjeturar una pendiente inicial α_1 como aproximación de $y'(a)$. A continuación, se resuelve el problema de valor inicial (PVI) siguiente:

$$\frac{d^2y}{dx^2} + g(y) = 0, \quad y(a) = A, \quad y'(a) = \alpha_1.$$

Sea $y(x, \alpha_1)$ la solución del problema anterior. Entonces, si $|y(x, \alpha_1) - B|$ es menor que la precisión deseada, entonces el proceso termina y $y(x, \alpha_1)$ es la solución del PVF (2.46). Pero lo más normal es no acertar a la primera con la aproximación de $y'(a)$. Por ello, se puede buscar otra aproximación α_2 para $y'(a)$ y calcular la solución $y(x, \alpha_2)$ del correspondiente PVI. Si $|y(x, \alpha_2) - B|$ sigue siendo mayor que la precisión deseada, se pueden usar diversos métodos iterativos (bisección, secante, Newton) para, a partir de α_1 y α_2, calcular nuevas aproximaciones $\alpha_3, \alpha_4, \dots$ de $y'(a)$. De esta forma, hay que resolver un PVI en cada paso.

Existen otras técnicas para resolver este tipo de problemas, como las basadas en el cálculo de variaciones. En el siguiente ejemplo, mostramos otra alternativa, propuesta por D. A. Sánchez [131].

Ejemplo 2.15. *El método de Newton se puede usar en una alternativa al método de disparo para resolver el PVF (2.46).*

La clave de esta alternativa está en observar que un PVI

$$\frac{d^2y}{dx^2} + g(y) = 0, \quad y(a) = A, \quad y'(a) = \alpha$$

satisface la ecuación de la energía

$$\frac{1}{2}\left(\frac{dy}{dx}\right)^2 + G(y) = \frac{\alpha^2}{2} + G(A), \quad G'(y) = g(y). \tag{2.47}$$

En efecto, derivando en (2.47),

$$y'y'' + G'(y)y' = 0 \Rightarrow y'' + g(y) = 0.$$

Además, sustituyendo $x = a$ en (2.47), se tiene $y'(a) = \alpha$.

Observemos ahora que, por el teorema de la función inversa, (2.47) puede escribirse

$$\left(\frac{dx}{dy}\right)^2 = \frac{1}{\alpha^2 + 2G(A) - 2G(y)}.$$

Entonces, si $A \neq B$, podemos encontrar la solución de esta ecuación:

$$x_\alpha = \pm \int_A^B \frac{1}{\sqrt{\alpha^2 + 2G(A) - 2G(y)}}\, dy.$$

El signo $+$ se considera si $B > A$ y el $-$ cuando $B < A$. A partir de ahora supondremos que $B > A$. El otro caso se estudia de forma parecida. La cantidad x_α se puede entender como el «tiempo» que emplea la solución para viajar del punto (A, α) a la recta vertical $y = B$ en el plano (y, y').

Después de estas consideraciones, el valor α^* de la pendiente que nos da la solución del PVF (2.46) es la solución de la ecuación no lineal $x_\alpha = b - a$, o equivalentemente,

$$F(\alpha) = b - a - \int_A^B \frac{1}{\sqrt{\alpha^2 + 2G(A) - 2G(y)}}\, dy = 0. \tag{2.48}$$

La existencia de una solución α^* de la ecuación (2.48) depende de la convergencia de la integral y, por tanto, de la función G. Bajo condiciones adecuadas, por ejemplo cuando G es una función monótonamente creciente o decreciente [131], se puede probar que la sucesión generada por el método de Newton:

$$\alpha_{n+1} = \alpha_n - \frac{F(\alpha_n)}{F'(\alpha_n)} = \alpha_n - \frac{b - a - \int_A^B (\alpha_n^2 + 2G(A) - 2G(y))^{-1/2}\, dy}{\alpha_n \int_A^B (\alpha_n^2 + 2G(A) - 2G(y))^{-3/2}\, dy}$$

converge a α^*.

En la práctica se puede tomar como punto de partida de esta iteración el valor $\alpha_1 = (B - A)/(b - a)$.

Capítulo 3

Dinámica del método de Newton en la recta real

3.1. Introducción

En el capítulo 1 ya se puso de manifiesto que el comportamiento dinámico de algunas funciones, aparentemente sencillas, como los polinomios de segundo grado o la función logística, puede ser extremadamente complicado. En esta sección nos vamos a centrar en estudiar el comportamiento dinámico del método de Newton, introducido en el capítulo 2, cuando se aplica para resolver una ecuación

$$f(x) = 0, \tag{3.1}$$

siendo f una función real de variable real. Dejamos para el capítulo 4 el análisis para funciones de variable compleja.

Así, sea $f : \mathbb{R} \to \mathbb{R}$ la función definida en la ecuación (3.1). La función de iteración para el método de Newton es entonces

$$N_f(x) = x - \frac{f(x)}{f'(x)}. \tag{3.2}$$

Nuestro objetivo es el estudio del comportamiento de las sucesiones $\{x_n\}$, donde

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = N_f(x_n) = N_f^n(x_0), \quad n \geq 0, \tag{3.3}$$

para distintos puntos de partida $x_0 \in \mathbb{R}$.

Una rápida inspección a la función de iteración del método de Newton nos permite sacar las siguientes conclusiones:

1) Si α es una raíz simple de la ecuación (3.1), es decir, satisface $f'(\alpha) \neq 0$, entonces α es un punto fijo de N_f, es decir, $N_f(\alpha) = \alpha$. Además, como

$$N_f'(\alpha) = \frac{f(\alpha)f''(\alpha)}{f'(\alpha)^2} = 0,$$

se trata de un punto fijo superatractor de la aplicación N_f. Notemos que este aspecto implica la convergencia cuadrática del método de Newton si comenzamos con un punto inicial cerca de la raíz buscada, tal y como se puso de manifiesto en la sección 2.4.

2) Cuando α es una raíz múltiple de la ecuación (3.1), entonces α también es un punto fijo de N_f, aunque en este caso se trata de un punto fijo atractor con

$$N_f'(\alpha) = \frac{m-1}{m} < 1,$$

siendo m la multiplicidad de α como raíz de f. Como vimos en la sección 2.5, la convergencia del método de Newton en estas condiciones es solamente lineal.

Un análisis un poco más profundo nos permite afirmar que, dada una ecuación (3.1), el comportamiento dinámico de la función de iteración del método de Newton (3.2) nos permite clasificar a los puntos de la recta real en dos conjuntos:

$$\mathcal{B}(f) = \{x_0 \in \mathbb{R} : \lim_{n\to\infty} N_f^n(x_0) = \alpha \quad \text{con} \quad f(\alpha) = 0\}, \tag{3.4}$$

$$\mathcal{M}(f) = \mathbb{R} - \mathcal{B}(f). \tag{3.5}$$

Coloquialmente hablando, podríamos definir a estos conjuntos como los puntos «buenos» y «malos» para el método de Newton con el siguiente sentido: $\mathcal{B}(f)$ recoge a los puntos iniciales x_0 a partir de los cuales el método de Newton (3.2) converge a una solución de la ecuación (3.1). Por tanto, en $\mathcal{M}(f)$ se incluyen los puntos de partida para los cuales, por un motivo u otro, el método de Newton no converge a ninguna solución de la ecuación considerada. Obsérvese que entre los puntos que forman parte del conjunto $\mathcal{M}(f)$ se encuentran los siguientes:

(a) Los puntos críticos de f, es decir, puntos para los cuales $f'(x_0) = 0$. Notemos que geométricamente estos puntos se caracterizan porque la tangente a la gráfica de f es horizontal.

(b) Las preimágenes de los puntos críticos, es decir, puntos x_0 tales que $f'(x_n) = 0$, siendo $x_n = N_f^n(x_0)$.

(c) Los n-ciclos de período $n \geq 2$ de N_f, así como los puntos que son atraídos por dichos ciclos.

Así, podemos precisar un poco más a la hora de clasificar los puntos contenidos en el conjunto $\mathcal{M}(f)$. Denotamos

$$\mathcal{A}(f) = \{x \in \mathbb{R} : \text{existe } m = m(x) \in \mathbb{N} \text{ tal que } N_f^m(x) \in Z(f')\}, \tag{3.6}$$

donde $Z(g) = \{x \in \mathbb{R} : g(x) = 0\}$ es el conjunto de los ceros de la función g. Notemos que $\mathcal{A}(f)$ está formado por las preimágenes de los puntos críticos de f:

$$\mathcal{A}(f) = \bigcup_{m \in \mathbb{N}} (N_f)^{-m}(Z(f')).$$

Toda vez que la función f tiene un cero t los puntos críticos de f son aislados, se cumple que $\mathcal{A}(f)$ es numerable y no denso en $\mathbb{R}$.

Englobamos dentro del conjunto

$$\mathcal{C}(f) = \mathcal{M}(f) - \mathcal{A}(f) \tag{3.7}$$

a los puntos con la dinámica más complicada respecto a la función de iteración N_f.

Los siguientes ejemplos nos muestran algunas de las formas que puede tomar el conjunto $\mathcal{M}(f)$ para distintas funciones f.

Ejemplo 3.1. *Si $f(x) = x^2 - 1$ entonces $\mathcal{M}(f)$ tiene un sólo un punto: $\mathcal{M}(f) = \{0\}$.*

En efecto, como veremos con más detalle en la siguiente sección, en este caso se tiene que si se parte de $x_0 > 1$ se obtiene una sucesión monótonamente decreciente a la raíz $\alpha = 1$. Si se parte de $x_0 \in (0, 1)$, entonces, $x_1 > 1$ y a partir de aquí se obtiene una sucesión monótonamente decreciente a la misma raíz anterior. El comportamiento para puntos de partida negativos es simétrico, obteniéndose en este caso sucesiones convergentes hacia la otra raíz, $\alpha = -1$. En consecuencia, el único punto de partida para el cual no existen las iteraciones del método de Newton es $x_0 = 0$.

Ejemplo 3.2. *Si $f(x) = x^3 - 1$ entonces $\mathcal{M}(f)$ tiene infinitos puntos. No obstante, el conjunto $\mathcal{M}(f)$ tiene medida cero.*

En este caso, la función de iteración del método de Newton es de la forma

$$N_f(x) = \frac{2x^3 + 1}{3x^2},$$

que no está definida en el punto $x = 0$ y en el conjunto de sus preimágenes (un conjunto numerable de puntos).

Ejemplo 3.3. *Si $f(x) = 25x^3 - 16x - 9$ entonces $\mathcal{M}(f)$ tiene infinitos puntos. Además, en este caso $\mathcal{M}(f)$ es un conjunto de medida positiva.*

En este caso, la función de iteración del método de Newton $N_f(x)$ tiene un 2-ciclo atractor de la forma

$$\{0.011036, -0.562826\}.$$

Es más, la función $g(x) = N_f^2(x)$ tiene un único punto fijo en el intervalo $[-0.025, 0.025]$ y satisface

$$|g'(x)| < 1, \quad x \in [-0.025, 0.025].$$

Por lo tanto las órbitas de cualquier punto de partida $x_0 \in [-0.025, 0.025]$ se verán atraídas hacia el 2-ciclo anterior. En consecuencia

$$[-0.025, 0.025] \subset \mathcal{M}(f).$$

Ejemplo 3.4. *El método de Newton aplicado al polinomio $p(x) = x^3 - 1.265x + 1$ tiene una 2-ciclo atractor.*

La función de iteración asociada es

$$N_p(x) = \frac{2x^3 - 1}{3x^2 - 1.265}.$$

En una primera inspección, calculamos la órbita del punto $x_0 = 0$:

$$\text{orb}(x_0) = \{0, 0.79051, -0.019675, 0.79125, -0.015043, 0.79094,$$

$$-0.016973, 0.79106, -0.016232, 0.79101\ \}.$$

Esta prueba numérica nos hace intuir la presencia de un 2-ciclo atractor. Para demostrar su existencia más rigurosamente, consideremos el intervalo $I = [-0.03, 0.03]$. Entonces se puede probar que:

(a) $N_p(I) \cap I = \emptyset$,

(b) $N_p^2 : I \to I$, y

(c) $|(N_p^2)'(x)| < 1$ para todo $x \in I$.

Las condiciones (a) y (b) junto con el teorema del valor intermedio garantizan la existencia de un punto fijo de N_p^2 que no es un punto fijo de N_p. Además, por (c), dicho punto fijo es un atractor.

Ejemplo 3.5. *Determínense algunos puntos de partida para los cuales el método de Newton aplicado a la función $f(x) = 2x^3 - 2x + \sqrt{2}$ no converja a ninguna raíz.*

En este caso $N_f(x) = (4x^3 - \sqrt{2})/(2(3x^2 - 1))$. Tenemos que $f'(x) = 0$ si y sólo si $x = \pm 1/\sqrt{3}$. En estos puntos la recta tangente al gráfico de f es horizontal.

Por otra parte, notemos que $N_f(0) = \sqrt{2}/2$ y $N_f(\sqrt{2}/2) = 0$. Por lo tanto, $\left\{0, \sqrt{2}/2\right\}$ es una órbita periódica de período 2 para N_f. Además, como

$$\begin{aligned} \left(N_f^2\right)'(x) &= N_f'\left(N_f(x)\right) \cdot N_f'(x) \\ &= \frac{f(N_f(x))f''(N_f(x))}{(f'(N_f(x)))^2} \frac{f(x)f''(x)}{(f'(x))^2}, \end{aligned}$$

y $f''(0) = 0$, se tiene que

$$(N_f^2)'(0) = \frac{f(\sqrt{2}/2)f''(\sqrt{2}/2)}{(f'(\sqrt{2}/2))^2}\frac{f(0)f''(0)}{(f'(0))^2} = 0,$$

por lo que $\left\{0, \sqrt{2}/2\right\}$ es una órbita periódica superatractora de N_f. En consecuencia, existe un intervalo I que contiene a 0 tal que si $x \in I$, entonces, los iterados pares $N_f^{2n}(x) \to 0$ y los iterados impares $N_f^{2n+1}(x) \to \sqrt{2}/2$.

Ejemplo 3.6. *Determínense algunos puntos de partida para los cuales el método de Newton aplicado a la función $f(x) = x^3 - x$ no converja a ninguna raíz.*

Las raíces de la ecuación $f(x) = 0$, son $0, 1$ y -1. La función de iteración de Newton para f es $N_f(x) = 2x^3/(3x^2 - 1)$. En este caso, tenemos que $f'(x) = 0$ si y sólo si $x = \pm\sqrt{3}/3$. Para estos valores de x el método de Newton no está definido.

Para ver si N_f tiene órbitas periódicas, por simetría, busquemos un punto x tal que $N_f(x) = -x$, es decir,

$$\frac{2x^3}{3x^2 - 1} = -x.$$

Al resolver esta ecuación obtenemos $x = 0$ o $x = \pm 1/\sqrt{5}$. El punto $x_0 = 0$ es un punto fijo (superatractor) de N_f, y los puntos $x_1 = 1/\sqrt{5}$ y $x_2 = -1/\sqrt{5}$ forman una órbita periódica de período 2, la cual es repulsora, pues $(N_f^2)'\left(1/\sqrt{5}\right) = 36$.

Estos ejemplos nos permiten intuir que el estudio de la convergencia global del método de Newton es un problema complicado, no sólo por la existencia de puntos donde la convergencia puede fallar, sino también porque en el caso en el que el método de Newton converge, puede hacerlo a una raíz distinta a la esperada. No siempre ocurre que el método de Newton converja a la raíz más próxima al punto de partida. Así, si en el último ejemplo tomamos $x_0 = -0.5$, se tiene que $N_f(-0.5) = 1$. Además, si tomamos puntos de partida cercanos a $-0, 5$, el método de Newton no converge a ninguna de las raíces más próximas, 0 ó -1, si no que lo hace a la raíz más alejada, $\alpha = 1$.

En el resto del capítulo nos centraremos en el estudio del comportamiento dinámico de la función de iteración del método de Newton (3.2) definida sobre la recta real. Así, en las secciones 3.2 y 3.3 analizamos el comportamiento particular de $N_f(x)$ para polinomios de segundo y tercer grado respectivamente, dejando para el resto de secciones el caso general.

3.2. Método de Newton para ecuaciones cuadráticas

Analizamos en esta sección del método de Newton para aproximar soluciones de las ecuaciones más simples después de las ecuaciones lineales, a saber, las ecuaciones cuadráticas.

Notemos que, sin pérdida de generalidad, podemos suponer que tenemos una ecuación de la forma

$$f(x) = x^2 + Bx + C = 0, \quad B, C \in \mathbb{R}, \tag{3.8}$$

ya que toda ecuación del tipo $\tilde{f}(x) = ax^2+bx+c$ con $a \neq 0$ se puede reducir a una del tipo (3.8) dividiendo por a y denotando $B = b/a$ y $C = c/a$. Nótese que además el comportamiento del método de Newton es el mismo para ambas funciones, es decir, $N_f(x) = N_{\tilde{f}}(x)$.

Es más, toda función cuadrática del tipo (3.8) puede reducirse a una del tipo

$$f_\mu(x) = x^2 - \mu, \quad \mu \in \mathbb{R} \tag{3.9}$$

mediante un cambio de coordenadas afín, como se puso de manifiesto en el ejemplo 1.13.

Como aplicación de este resultado, podemos reducir el estudio del comportamiento dinámico del método de Newton aplicado a polinomios de segundo grado a los polinomios de la familia uniparamétrica (3.9).

Teorema 3.1. *Sean $f(x)$ y $f_\mu(x) = x^2 - \mu$, definidos en (3.8) y (3.9) respectivamente, con $\mu = (B^2 - 4C)/4$. Entonces $N_f(x)$ es conjugada topológicamente con $N_{f_\mu}(x)$ mediante la transformación afín $h(x) = x - \frac{B}{2}$, es decir, $h^{-1} \circ N_f \circ h(x) = N_{f_\mu}(x)$.*

Demostración. En primer lugar, es una comprobación inmediata que $f_\mu(x) = f \circ h(x)$. Además, como $h'(x) = 1$, se sigue que $f'(h(x)) = (f \circ h)'(x) = f'_\mu(x)$. Así,

$$\begin{aligned} h^{-1} \circ N_f \circ h(x) &= h^{-1}\left(h(x) - \frac{f(h(x))}{f'(h(x))}\right) \\ &= h(x) - \frac{f(h(x))}{f'(h(x))} + \frac{B}{2} \\ &= x - \frac{f_\mu(x)}{f'_\mu(x)} \\ &= N_{f_\mu}(x). \end{aligned}$$

□

El siguiente paso es caracterizar, en función del signo de μ, el comportamiento del método de Newton aplicado a un polinomio de la forma (3.9). En concreto, el siguiente resultado muestra que el estudio de las iteraciones dadas por el método de Newton aplicado a un polinomio cuadrático general, se puede reducir a tres situaciones particulares, dependiendo de si el polinomio original tiene dos raíces reales, ninguna raíz real o una raíz doble.

Teorema 3.2. *Sea N_μ la función de iteración del método de Newton aplicado a un polinomio de la forma (3.9). Entonces*

1. *Si* $\mu > 0$, N_μ *es conjugado topológicamente con la aplicación* $N_1(x)$, *donde*
$$N_1(x) = \frac{x^2+1}{2x}$$
es el método de Newton aplicado al polinomio $f_1(x) = x^2 - 1$.

2. *Si* $\mu < 0$, *entonces* N_μ *es conjugado topológicamente con la aplicación*
$$N_{-1}(x) = \frac{x^2-1}{2x},$$
que es el método de Newton aplicado al polinomio $f_{-1}(x) = x^2 + 1$.

3. *Si* $\mu = 0$, *nos queda la aplicación lineal* $N_0(x) = \frac{x}{2}$, *que es el método de Newton aplicado al polinomio* $f_0(x) = x^2$.

Demostración. En el primer caso, busquemos una transformación afín $H(x) = \alpha x + \beta$, de modo que tengamos la ecuación de conjugación
$$H \circ N_\mu = N_1 \circ H. \tag{3.10}$$
Ahora, $H \circ N_\mu(x) = \alpha N_\mu(x) + \beta$ y $N_1 \circ H(x) = (H(x)^2 + 1)/(2H(x))$. Usando la ecuación de conjugación (3.10), tenemos la ecuación
$$(\alpha x^2 + 2\beta x + \alpha\mu)(\alpha x + \beta) = x(\alpha^2 x^2 + 2\alpha\beta x + \beta^2 + 1)$$
que, después de realizar todas las simplificaciones posibles, nos queda como
$$3\alpha\beta x + 2\beta^2 + \alpha^2\mu = 2\alpha^2\beta + \beta^2 + 1.$$
De aquí, $\alpha\beta = 0$ y como $\alpha \neq 0$ se sigue que $\beta = 0$. También como $2\beta^2 + \alpha^2\mu = 2\alpha^2\beta + \beta^2 + 1$, nos queda $\alpha^2\mu = 1$, de donde $\alpha = 1/\sqrt{\mu}$, concluyendo que la transformación afín buscada es $H(x) = x/\sqrt{\mu}$.

En el segundo caso, como en el caso anterior, buscamos $H(x) = \alpha x + \beta$ afín, tal que $H \circ N_\mu = N_{-1} \circ H$, lo que nos conduce a la ecuación $\alpha^2\mu = -1$, de donde $\alpha = 1/\sqrt{-\mu}$, y $H(x) = x/\sqrt{-\mu}$ es la transformación buscada.

El tercer caso es una comprobación inmediata. □

Corolario 3.3. *El comportamiento dinámico del método de Newton aplicado a los polinomios de la forma (3.8) se reduce a tres casos particulares, en función del signo de* $\Delta = B^2 - 4C$, *el discriminante de la correspondiente ecuación cuadrática. En concreto,*

1. *Si* $\Delta = 0$, N_f *es conjugada topológicamente con* $N_0(x)$, *el método de Newton aplicado al polinomio* $f_0(x) = x^2$.

2. *Si* $\Delta > 0$, N_f *es conjugada topológicamente con* $N_1(x)$, *el método de Newton aplicado al polinomio* $f_1(x) = x^2 - 1$.

3. *Si* $\Delta < 0$, *entonces* N_f *es conjugada topológicamente con* $N_{-1}(x)$, *el método de Newton aplicado al polinomio* $f_{-1}(x) = x^2 + 1$.

El método de Newton aplicado al polinomio $f_0(x) = x^2$

Esta situación es la más fácil de las tres, ya que la función de iteración del método de Newton queda de la forma

$$N_0(x) = \frac{x}{2}. \tag{3.11}$$

En consecuencia, la sucesión generada por dicha función es $x_n = x_0/2^n$, por lo que $x_n \to 0$ para todo $x_0 \in \mathbb{R}$. Nótese que, en este caso, la convergencia es lineal.

El método de Newton aplicado al polinomio $f_1(x) = x^2 - 1$

En este caso la función de iteración del método de Newton es

$$N_1(x) = \frac{x^2+1}{2x} = \frac{1}{2}\left(x + \frac{1}{x}\right). \tag{3.12}$$

El único punto donde no está definida es en $x = 0$. Es sencillo comprobar que si $x_0 > 1$, entonces la sucesión $x_{n+1} = N_1(x_n)$ es monótonamente decreciente hacia 1, es decir, $1 < x_{n+1} < x_n, \quad n \geq 0$.

Si $0 < x_0 < 1$, entonces $x_1 > 1$ y a partir de aquí se inicia la convergencia monótona decreciente hacia 1. En este caso, $x_0 < x_1$ y $1 < x_{n+1} < x_n, \quad n \geq 1$.

El estudio para valores negativos de x_0 se puede hacer por simetría, ya que si $y_0 = -x_0$, con $x_0 > 0$, entonces si $y_{n+1} = N_1(y_n)$ y $x_{n+1} = N_1(x_n)$, se tiene que $y_n = -x_n$. Por lo tanto, si $y_0 < -1$, entonces la sucesión $\{y_n\}$ es monótonamente creciente a -1, mientras que si $-1 < y_0 < 0$, entonces $y_1 < -1$ y a partir de aquí, la sucesión crece hacia el límite -1.

El método de Newton aplicado al polinomio $f_1(x) = x^2 + 1$

En este caso se comprueba que la función de iteración obtenida aplicando el método de Newton al polinomio $f_1(x) = x^2 + 1$:

$$N_{-1}(x) = \frac{x^2-1}{2x}, \tag{3.13}$$

tiene un comportamiento caótico. En concreto, veamos que N_{-1} es conjugada topológicamente con la función «diente de sierra» definida en el ejemplo 1.16 de la siguiente manera:

$$S(x) = \begin{cases} 2x & \text{si} \quad 0 \leq x < \dfrac{1}{2}, \\ 2x - 1 & \text{si} \quad \dfrac{1}{2} \leq x \leq 1. \end{cases} \tag{3.14}$$

En dicho ejemplo ya se probó el comportamiento caótico de las órbitas de la función (3.14).

Consideramos la función $g : (0, 1) \to \mathbb{R}$ definida por

$$g(x) = -\frac{\cos(\pi x)}{\operatorname{sen}(\pi x)}.$$

Entonces, se tiene que $g \circ S(x) = N_{-1} \circ g(x)$ para todo $x \in (0, 1/2) \cup (1/2, 1)$. En efecto, por una parte se tiene que

$$N_{-1} \circ g(x) = -\frac{1 - \operatorname{tg}^2(\pi x)}{2\operatorname{tg}(\pi x)} = -\frac{1}{\operatorname{tg}(2\pi x)}.$$

Por otra parte, si $x \in (0, 1/2)$, entonces

$$g \circ S(x) = g(2x) = -\frac{1}{\operatorname{tg}(2\pi x)}.$$

Si $x \in (1/2, 1)$, entonces

$$g \circ S(x) = g(2x - 1) = -\frac{1}{\operatorname{tg}(2\pi x - \pi)} = -\frac{1}{\operatorname{tg}(2\pi x)}.$$

Por lo tanto, en ambos casos se tiene que $g \circ S(x) = N_{-1} \circ g(x)$.

Además, si tenemos en cuenta que

$$\lim_{x \to 0^+} g \circ S(x) = -\infty = \lim_{x \to 0^+} N_{-1} \circ g(x), \quad \lim_{x \to 1^-} g \circ S(x) = \infty = \lim_{x \to 1^-} N_{-1} \circ g(x),$$

la igualdad $g \circ S(x) = N_{-1} \circ g(x)$ se puede extender al intervalo $[0, 1]$, definiendo $g(0) = -\infty$; $g(1) = \infty$; $N_{-1}(-\infty) = -\infty$ y $N_{-1}(\infty) = \infty$.

Como el comportamiento del sistema dinámico (X, S), con $X = [0, 1]$ y S definida en (3.14), es caótico (véase el ejemplo 1.16), también lo es el de la función N_{-1} sobre toda la recta real.

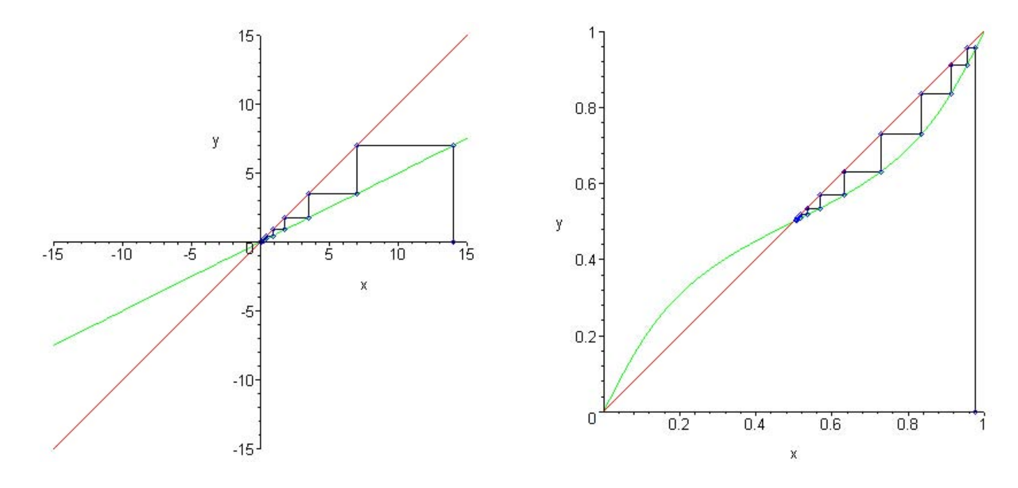


Figura 3.1: Gráfico de N_0 en $\mathbb{R}$ y gráfico de su compactificación $\tilde{N}_0$ en el intervalo $[0, 1]$.

Para terminar esta sección, a modo de resumen, mostramos los gráficos de algunas iteraciones del método de Newton aplicado a los tres polinomios a los que se reduce el estudio para polinomios cuadráticos, junto con sus correspondientes compactificaciones al intervalo $[0, 1]$, entendiendo por compactificación de una función $f(x)$ definida en la recta real como otra

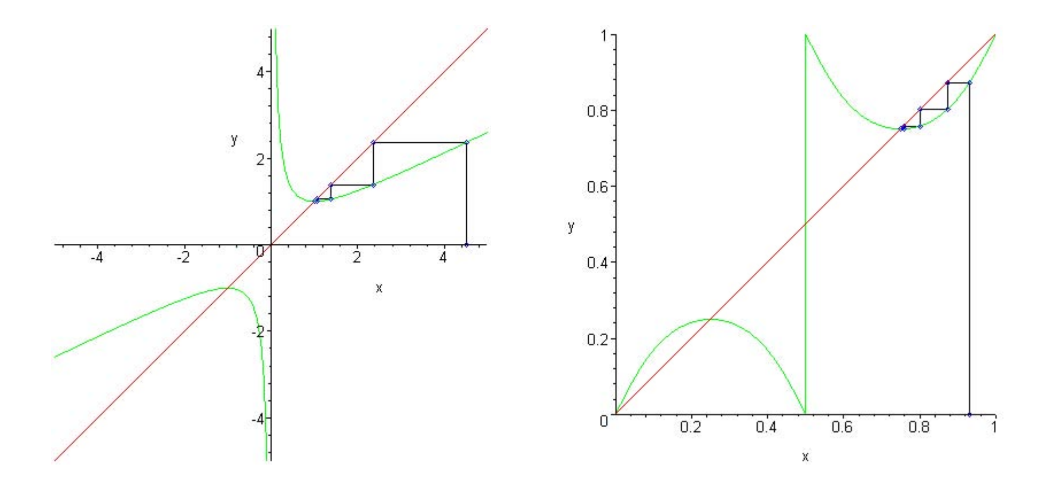


Figura 3.2: Gráfico de N_1 en $\mathbb{R}$ y gráfico de su compactificación $\tilde{N}_1$ en el intervalo $[0, 1]$.

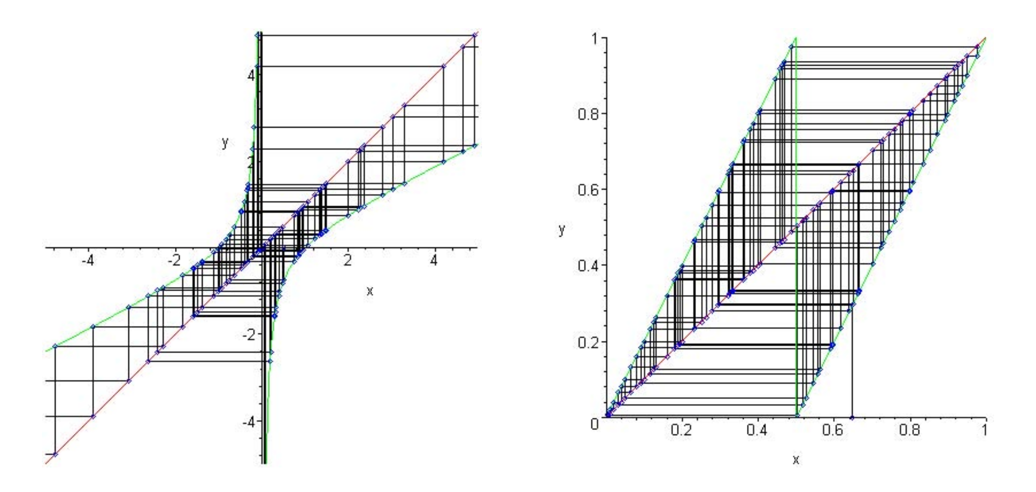


Figura 3.3: Gráfico de N_{-1} en $\mathbb{R}$ y gráfico de su compactificación $\tilde{N}_{-1}$ en el intervalo $[0, 1]$.

función que es conjugada topológicamente con $f(x)$, pero que está definida en un intervalo finito.

Para obtener dichas compactificaciones, se considera la aplicación $G : \mathbb{R} \to (0, 1)$ dada por

$$G(x) = \frac{1}{\pi} \operatorname{arc\,tg}(x) + \frac{1}{2}.$$

Esta función es un homeomorfismo de $\mathbb{R}$ en el intervalo $(0, 1)$. Por lo tanto, para las funciones N_0, N_1 y N_{-1} introducidas en (3.11), (3.12) y (3.13) respectivamente, construimos las compactificaciones

- $\tilde{N}_0 : (0, 1) \to (0, 1)$ definida por $\tilde{N}_0(x) = (G \circ N_0 \circ G^{-1})(x)$;
- $\tilde{N}_1 : (0, 1) \to (0, 1)$ definida por $\tilde{N}_1(x) = (G \circ N_1 \circ G^{-1})(x)$;
- $\tilde{N}_{-1} : (0, 1) \to (0, 1)$ definida por $\tilde{N}_{-1}(x) = (G \circ N_{-1} \circ G^{-1})(x)$.

Además, si introducimos las notaciones

$$G(-\infty) = \lim_{x\to-\infty} G(x) = 0, \quad G(+\infty) = \lim_{x\to+\infty} G(x) = 1,$$

y tenemos en cuenta que

$$\lim_{x\to\infty} N_0(x) = \lim_{x\to\infty} N_1(x) = \lim_{x\to\infty} N_{-1}(x) = \infty,$$

$$\lim_{x\to-\infty} N_0(x) = \lim_{x\to-\infty} N_1(x) = \lim_{x\to-\infty} N_{-1}(x) = -\infty,$$

podemos extender cualquiera de las tres compactificaciones a un aplicación definida en $[0,1]$. Usamos la misma notación para esta extensión. Notemos que las funciones extendidas tienen a $x = 0$ y a $x = 1$ como puntos fijos adicionales a los puntos fijos que corresponden a las raíces del polinomio en cuestión.

3.3. Método de Newton para polinomios cúbicos

En la sección anterior ya observamos que el método de Newton puede presentar un comportamiento caótico para algunos polinomios cuadráticos. Ahora bien, para polinomios cúbicos, el comportamiento de la transformada de Newton es más interesante desde el punto dinámico, con la aparición de fenómenos interesantes como las cascadas de bifurcaciones.

En esta sección analizamos la dinámica de la función de iteración del método de Newton

$$N_p(x) = x - \frac{p(x)}{p'(x)},$$

donde $p(x)$ es un polinomio cúbico de la forma

$$p(x) = x^3 + bx^2 + cx + d, \quad b, c, d \in \mathbb{R}. \tag{3.15}$$

Podemos suponer, sin pérdida de generalidad que $p(x)$ es mónico ya que la función de iteración del método de Newton es la misma para los polinomios $Ax^3 + Bx^2 + Cx + D$ $(A \neq 0)$ y $x^3 + bx^2 + cx + d$ con $b = B/A$, $c = C/A$ y $d = D/A$.

El número de parámetros implicados en el problema se puede reducir a dos, teniendo en cuenta el siguiente resultado.

Lema 3.4. *Sea $p(x)$ un polinomio de la forma (3.15) y sea $\tau(x) = 3x + b$. Entonces N_p es conjugada topológicamente con N_q, la función de iteración del método de Newton aplicada al polinomio*

$$q(x) = x^3 + \lambda x + \gamma, \tag{3.16}$$

donde $\lambda = 9c - 3b^2$ y $\gamma = 2b^3 - 9(bc - 3d) = 27d + 2b^3 - 9bc$.

Demostración. En primer lugar, notemos que

$$N_p(x) = \frac{2x^3 + bx^2 - d}{3x^2 + 2bx + c} \quad \text{y} \quad N_q(x) = \frac{2x^3 - C}{3x^2 + B}.$$

Entonces

$$\tau \circ N_p(x) = \frac{6x^3 + 6bx^2 + 2b^2x + bc - 3d}{3x^2 + 2bx + c}$$

y

$$N_q \circ \tau(x) = \frac{54x^3 + 54bx^2 + 18b^2x + 2b^3 - \gamma}{27x^2 + 18bx + 3b^2 + \lambda}.$$

Teniendo en cuenta los valores de λ y γ dados en el enunciado del teorema, se sigue que $\tau \circ N_p(x) = N_q \circ \tau(x)$, es decir, N_p y N_q son conjugadas topológicamente. □

Pero aún se puede reducir más el estudio del método de Newton para polinomios cúbicos, sin necesidad de considerar todos los polinomios definidos en (3.16). En concreto, el siguiente resultado prueba que basta con estudiar el método de Newton en cuatro situaciones concretas.

Teorema 3.5. *Sea N_q la función de iteración del método de Newton aplicado a un polinomio de la forma (3.16), es decir,*

$$N_q(x) = \frac{2x^3 - \gamma}{3x^2 + \lambda}.$$

Entonces

1. *Si $\lambda = \gamma = 0$, entonces $N_q(x) = N_0(x) = 2x/3$, que es el método de Newton aplicado al polinomio $p_0(x) = x^3$, que es un polinomio con una raíz triple en $x = 0$.*

2. *Si $\gamma = 0$ y $\lambda > 0$, entonces $N_q(x)$ es conjugada topológicamente con la aplicación $N_+(x)$, donde*

$$N_+(x) = \frac{2x^3}{3x^2 + 1}$$

es el método de Newton aplicado al polinomio $p_+(x) = x^3 + x$, que es un polinomio con una raíz real y dos complejas conjugadas.

3. *Si $\gamma = 0$ y $\lambda < 0$, entonces $N_q(x)$ es conjugada topológicamente con la aplicación $N_-(x)$, donde*

$$N_-(x) = \frac{2x^3}{3x^2 - 1}$$

es el método de Newton aplicado al polinomio $p_-(x) = x^3 - x$, que es un polinomio con tres raíces reales diferentes.

4. *Si $\gamma \neq 0$, entonces $N_q(x)$ es conjugada topológicamente con la aplicación $N_c(x)$, donde*

$$N_c(x) = \frac{2x^3 - 1}{3x^2 + c} \tag{3.17}$$

es el método de Newton aplicado al polinomio $p_c(x) = x^3 + cx + 1$.

Demostración. El primer caso es una comprobación inmediata. Para ver el segundo caso, consideremos la transformación afín $h(x) = \sqrt{\lambda}x$. Entonces tenemos

$$\begin{aligned} h^{-1} \circ N_q \circ h(x) &= h^{-1}\left(\frac{2\lambda\sqrt{\lambda}x^3}{3\lambda x^2 + \lambda}\right) \\ &= \frac{1}{\sqrt{\lambda}}\frac{2\lambda\sqrt{\lambda}x^3}{3\lambda x^2 + \lambda} \\ &= \frac{2x^3}{3x^2 + 1} = N_+(x). \end{aligned}$$

El tercer caso es análogo, pero considerando la transformación afín $h(x) = \sqrt{|\lambda|}x$.

En el cuarto caso, escribimos el polinomio (3.15) de la forma $q(x) = x^3 + \lambda x + \beta^3$, es decir, con $\gamma = \beta^3$. Entonces, para $\tau(x) = x/\beta$ se tiene que $\tau \circ N_q = N_c \circ \tau$ con $p_c(x) = x^3 + cx + 1$ y $c = \lambda/\beta^2$. □

A continuación estudiamos con más detalle cada uno de los casos que se presentan en el teorema 3.5.

El método de Newton aplicado al polinomio $f_0(x) = x^3$

En este caso la función de iteración del método de Newton es $N_0(x) = 2x/3$. Por lo tanto, la sucesión generada por dicha función es $x_n = (2/3)^n x_0$, por lo que $x_n \to 0$ para todo $x_0 \in \mathbb{R}$. Nótese que, en este caso, la convergencia es lineal.

El método de Newton aplicado al polinomio $f_+(x) = x^3 + x$

En este caso, el único punto fijo de la correspondiente función de iteración del método de Newton

$$N_+(x) = \frac{2x^3}{3x^2 + 1}$$

es $x = 0$. Es sencillo comprobar que si $x_0 > 0$, entonces la sucesión $x_{n+1} = N_+(x_n)$ es monótonamente decreciente hacia 0, es decir $0 < x_{n+1} < x_n, \quad n \geq 0$. Teniendo en cuenta la simetría respecto al origen de la función N_+, se tiene que para puntos de partida $x_0 < 0$ se obtiene una sucesión que crece monótonamente hacia 0.

El método de Newton aplicado al polinomio $f_-(x) = x^3 - x$

Para las iteraciones de la función asociada a $f_-(x) = x^3 - x$,

$$N_-(x) = \frac{2x^3}{3x^2 - 1},$$

existen tres puntos fijos, que corresponden a las raíces $r_1 = -1$, $r_2 = 0$ y $r_3 = 1$. En este caso no es sencillo decidir hacia que raíz se aproximan las iteraciones de un punto arbitrario.

Como en el caso de las aplicaciones cuadráticas, consideramos la aplicación $G : \mathbb{R} \to (0,1)$ dada por

$$G(x) = \frac{1}{\pi} \operatorname{arc\,tg}(x) + \frac{1}{2}$$

y la correspondiente compactificación de la función N_{-}: $\tilde{N}_{-} : (0,1) \to (0,1)$ definida por

$$\tilde{N}_{-}(x) = (G \circ N_{-} \circ G^{-1})(x).$$

Notemos que $G(-1) = 1/4$, $G(0) = 1/2$ y $G(1) = 3/4$, y las asíntotas verticales están sobre los puntos $a_1 = -\sqrt{3}/3$ y $a_2 = \sqrt{3}/3$, que son llevadas por G en los puntos $\tilde{a}_1 = 1/3$ y $\tilde{a}_2 = 2/3$. La preimagen α por $\tilde{N}$ de $\tilde{a}_1 = 1/3$, distinta de este punto fijo, determina un intervalo $I_2 = (\alpha, \tilde{a}_2)$, de modo que las iteraciones por $\tilde{N}_{-}$ de puntos en ese intervalo convergen al punto fijo $1/4$, de modo análogo, tenemos un punto β, preimagen por $\tilde{N}_{-}$ de $2/3$, determina un intervalo $I_1 = (\tilde{a}_1, \beta)$, de modo que las iteraciones de puntos en ese intervalo por $\tilde{N}_{-}$ convergen al punto fijo $3/4$. Hay una cantidad numerable de puntos donde no se tiene convergencia correspondientes a la unión de las preimágenes por $\tilde{N}_{-}$ de las asíntotas verticales $\tilde{a}_1$ y $\tilde{a}_2$, es decir, el conjunto $\tilde{A} = \bigcup_{n\in\mathbb{N}} \tilde{N}_{-}^{-n}(\{\tilde{a}_1, \tilde{a}_2\})$, y existe un conjunto abierto constituido por la unión de las preimágenes de las cuencas de atracción de los puntos fijos $1/4$ y $3/4$, las cuales aparecen intercaladas, es decir, $\bigcup_{n\in\mathbb{N}}(\tilde{N}_{-}^{-n}(0, \tilde{a}_1) \cup \tilde{N}_{-}^{-n}(\tilde{a}_2, 1))$. Como se observa, decidir a cuál de los puntos fijos convergen las iteraciones, no es una tarea fácil en este caso. Para mayores detalles, el lector puede consultar [154].

El método de Newton aplicado al polinomio $f_c(x) = x^3 + cx + 1$

Esta situación esconde una gran riqueza desde el punto de vista dinámico. De hecho, podemos encontrar una cáscada de bifurcaciones de período, que recuerdan al diagrama de Feigenbaum visto para la función logística en el capítulo 1.

Para el polinomio $f_c(x)$, consideramos la correspondiente función de iteración del método de Newton (3.17). Como

$$N_c'(x) = \frac{f_c(x) f_c''(x)}{f_c'(x)^2} = \frac{6x(x^3 + cx + 1)}{3x^2 + c},$$

el único punto crítico libre de $N_c(x)$ es $x = 0$. Para obtener valores de c para los cuales $N_c(x)$ tiene un ciclo de la forma $0 \mapsto x_1 \mapsto x_2 \mapsto x_3 \mapsto \cdots \mapsto x_n \mapsto 0$, debemos resolver las ecuaciones

$$N_c^2(0) = N_c \circ N_c(0) = 0, \quad N_c^4(0) = N_c^2 \circ N_c^2(0) = 0, \quad N_c^8(0) = N_c^4 \circ N_c^4(0) = 0,$$

y así sucesivamente, de este modo obtenemos aquellos valores de c para los cuales N_c tiene un ciclo superatractor que contiene a $x = 0$, y por continuidad, obtenemos valores del parámetro

para los cuales $N_c(x)$ tiene un ciclo atractor del mismo período que el ciclo que contiene al punto 0. En concreto, la ecuación $N_c^2(0) = 0$ viene dada por

$$N_c^2(0) = -\frac{2 + c^3}{c(c^3 + 3)} = 0.$$

Dicha ecuación tiene como solución real $c_1 \approx -1.25992$. Así obtenemos el ciclo $x_0 = 0 \mapsto x_1 \approx 0.79370 \mapsto x_2 = x_0$.

La próxima ecuación a resolver es $N_c^4(0) = 0$, la cual tiene varias soluciones reales, la primera después de c_1 es $c_2 \approx -1.27816$ (mirando desde la derecha a la izquierda), para la cual N_{c_2} tiene un 4-ciclo:

$$x_0 = 0 \mapsto x_1 \approx 0.78237 \mapsto x_2 \approx -0.07561 \mapsto x_3 \approx 0.79370 \mapsto x_4 = x_0.$$

La próxima solución, correspondiente a $N_c^8(0) = 0$, es $c_3 \approx -1.29376$, para cual N_{c_3} tiene el 8-ciclo siguiente:

$$x_0 = 0 \mapsto x_1 \approx 0.77294 \mapsto x_2 \approx -0.15331 \mapsto x_3 \approx 0.82339 \mapsto x_4 \approx 0.15734 \mapsto$$

$$x_5 \approx 0.81363 \mapsto x_6 \approx 0.11156 \mapsto x_7 \approx 0.79370 \mapsto x_8 = x_0.$$

Las figuras 3.4 y 3.5 muestran los ciclos superatractores que hemos encontrado.

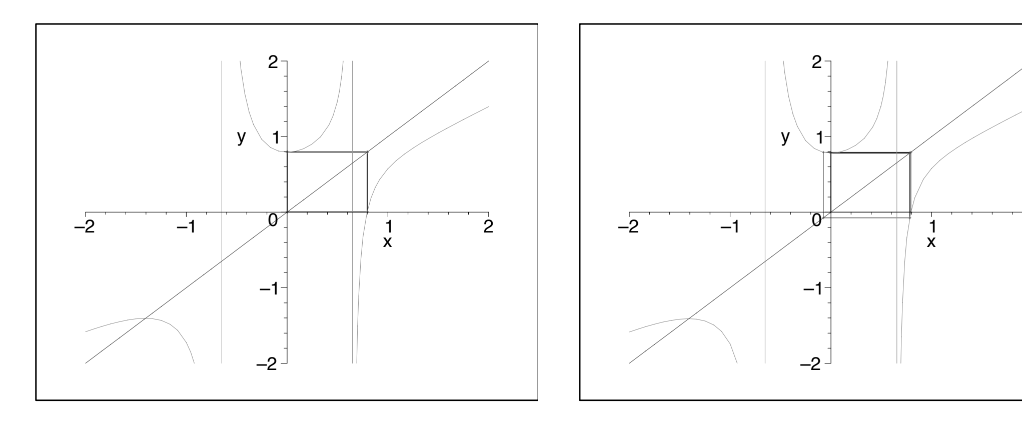


Figura 3.4: Órbita de 0 por N_{c_1} (un 2-ciclo) y órbita de 0 por N_{c_2} (un 4-ciclo).

Las ecuaciones para obtener ciclos de mayor período para N_c vienen dadas por funciones racionales en c, con numerador de un grado muy alto. El estudio dinámico de las funciones (3.17) da lugar a un diagrama de bifurcación como el que se muestra en la figura 3.6. Esta figura guarda una cierta similitud con el diagrama de Feighenbaum (véase la figura 1.11) para la ecuación logística y, al igual que él, muestra los valores del parámetro c para los cuales el método de Newton aplicado a un polinomio de la familia (3.17) tiene un punto fijo, a partir de cuándo aparece un 2-ciclo, un 4-ciclo, etc.

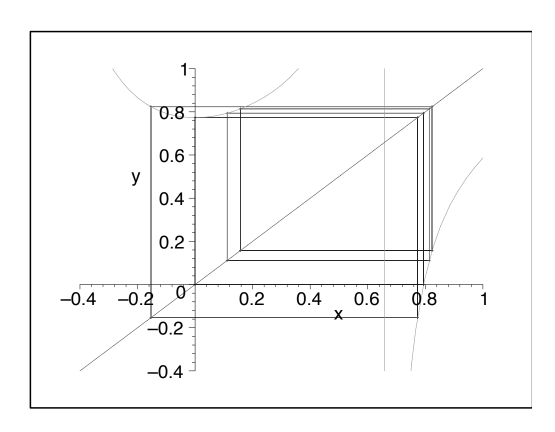


Figura 3.5: Órbita de 0 por N_{c_3} (un 8-ciclo).

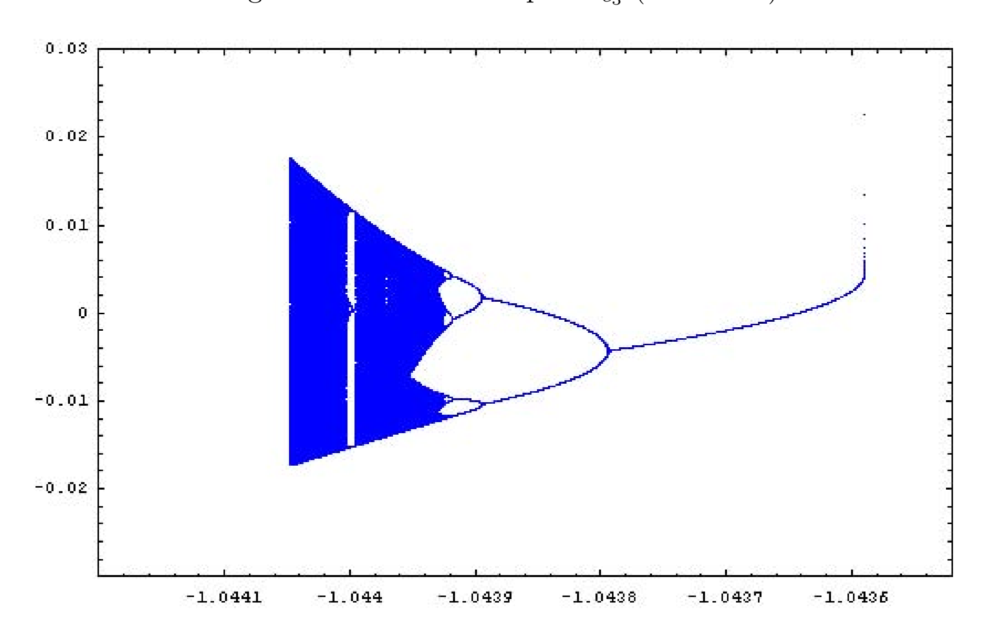


Figura 3.6: Diagrama de bifurcación para la familia uniparamétrica (3.17). En el eje de abcisas aparecen los valores del parámetro c y en el eje de ordenadas los límites de la correspondiente sucesión generada por el método de Newton, apreciándose la aparición de 2-ciclos, 4-ciclos, etc.

3.4. Propiedades básicas de la función de iteración del método de Newton

Como hemos comentado en la sección anterior, la propiedad fundamental de la función de iteración de Newton, N_f, de una cierta función f, es que transforma el problema de encontrar las raíces de la ecuación $f(x) = 0$ en el problema de encontrar los puntos fijos de N_f. En esta sección vamos a realizar un estudio más detallado de la función de iteración del método de Newton (3.2) para funciones $f : \mathbb{R} \to \mathbb{R}$ suficientemente derivables. Para simplificar nuestro estudio, supongamos que f satisface las condiciones siguientes:

C.1 Si $f'(x) = 0$, entonces $f''(x) \neq 0$, es decir, los puntos críticos (máximos o mínimos) de

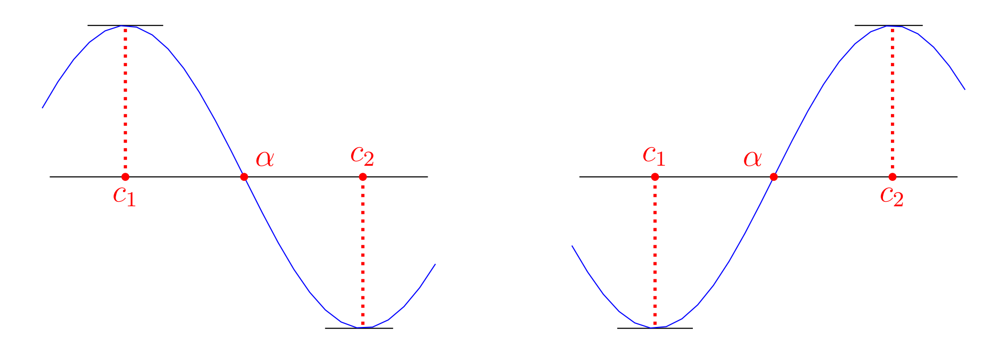


Figura 3.7: Posibles gráficos de $f(x)$ en una banda (c_1, c_2) que contiene una raíz α de $f(x) = 0$.

f son no degenerados.

C.2 Si $f(x) = 0$, entonces $f'(x) \neq 0$. Es decir, f sólo tiene raíces simples. Geométricamente si $f(\alpha) = 0$ y $f'(\alpha) \neq 0$ entonces f es transversal al eje OX en el punto $(\alpha, 0)$, o equivalentemente el gráfico de f cruza en el punto $(\alpha, 0)$ el eje de las abscisas con un ángulo no cero.

Las siguientes propiedades son básicas para el estudio de las iteraciones de N_f.

(1) $f(\alpha) = 0$ si y sólo si $N_f(\alpha) = \alpha$. Además, si $f(\alpha) = 0$ entonces $N_f'(\alpha) = 0$, es decir, α es un punto fijo superatractor y $\lim_{k\to\infty} N_f^k(x) = \alpha$ para todo x suficientemente próximo a α.

(2) N_f tiene una asíntota vertical en cada solución real $x = c$ de $f'(x) = 0$.

Si $c_1 < c_2$ son dos raíces consecutivas de $f'(x) = 0$, el intervalo (c_1, c_2) se llama una *banda acotada* para N_f. Si c es la mayor de las raíces de $f'(x) = 0$, entonces el intervalo $(c, +\infty)$ es una *banda extrema* de N_f. Análogamente, si b es la menor de las raíces de $f'(x) = 0$, entonces el intervalo $(-\infty, b)$ es la otra *banda extrema* de N_f.

(3) Si (c_1, c_2) es una banda para N_f que contiene una raíz, α, de $f(x) = 0$, entonces

$$\lim_{x\to c_1^+} N_f(x) = +\infty, \qquad \lim_{x\to c_2^-} N_f(x) = -\infty.$$

En efecto, como $f'(x)$ tiene signo constante en (c_1, c_2),tenemos dos posibilidades para el gráfico local de f en el intervalo (c_1, c_2), que son las que se muestran en la figura 3.7.

En el primer caso, para $x \in (c_1, \alpha)$, se tiene que $f(x) > 0$ y $f'(x) < 0$. Luego

$$\lim_{x\to c_1^+} N_f(x) = \lim_{x\to c_1^+} x - \frac{f(x)}{f'(x)} = \infty.$$

En el intervalo (α, c_2), $f(x) < 0$ y $f'(x) < 0$. Luego

$$\lim_{x\to c_2^-} N_f(x) = \lim_{x\to c_2^-} x - \frac{f(x)}{f'(x)} = -\infty.$$

El segundo caso es completamente análogo.

(4) Si (c_1, c_2) es una banda de N_f, que no contiene raíces de $f(x) = 0$, entonces

$$\lim_{x\to c_1^+} N_f(x) = \lim_{x\to c_2^-} N_f(x) = \pm\infty.$$

En efecto, en este caso basta observar las posibles gráficas de f que se muestran en la figura 3.8 para deducir el resultado.

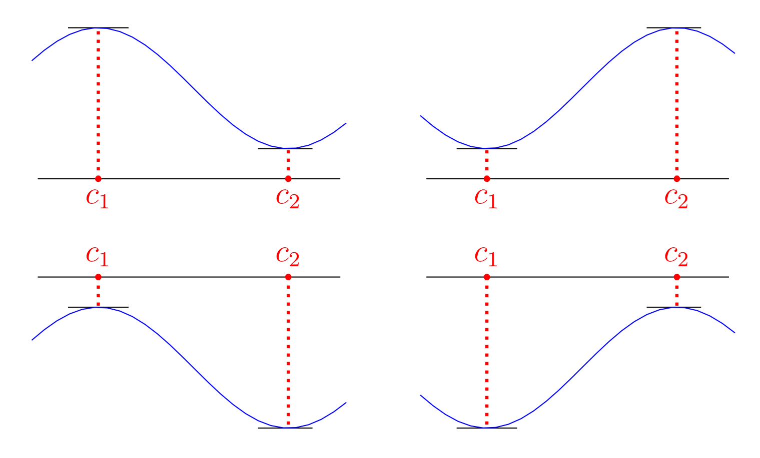


Figura 3.8: Posibles gráficos de $f(x)$ en una banda (c_1, c_2) que no contiene ninguna raíz α de $f(x) = 0$.

(5) Los puntos extremos locales de N_f, si $f'(x) \neq 0$, son los puntos para los cuales $f(x) = 0$ o $f''(x) = 0$, esto es, los ceros y los puntos de inflexión de f. En efecto, tenemos que $N_f'(x) = f(x)f''(x)/f'(x)^2 = 0$ si, y sólo si, $f(x) = 0$ o $f''(x) = 0$.

(6) Si una banda de N_f no contiene una raíz de $f(x) = 0$ entonces una de las bandas adyacentes o una banda extrema no contiene raíces de $f(x) = 0$. Esta propiedad frecuentemente vale también para bandas extremas. Una condición suficiente para que esta propiedad sea verdadera para bandas extremas es que $f''(x)$ sea siempre acotada desde 0 cuando $|x| \to +\infty$. Por ejemplo los polinomios tienen esta propiedad, sin embargo esta falla para $f(x) = xe^{-x}$. En este caso tenemos que $f(x) = 0$ si, y sólo si, $x = 0$. Además, $f'(x) = e^{-x}(1 - x)$. Luego las asíntotas de N_f, que son las soluciones de $f'(x) = 0$, se reducen a solamente $x = 1$. Pero entonces,

$$N_f(x) = x - \frac{xe^{-x}}{e^{-x}(1-x)} = -\frac{x^2}{1-x} = \frac{x^2}{x-1}.$$

En este caso, N_f tiene dos bandas: $(-\infty, 1)$ y $(1, \infty)$. En la banda $(1, \infty)$ no hay ninguna raíz de f y, sin embargo, sí que la hay en la otra banda. Por lo tanto, la función $f(x) = xe^{-x}$ no cumple la propiedad (6).

Finalmente de las propiedades (1)–(6), el gráfico típico de N_f, con f que satisface **C.1** y **C.2**, es como el que se muestra en la figura 3.9.

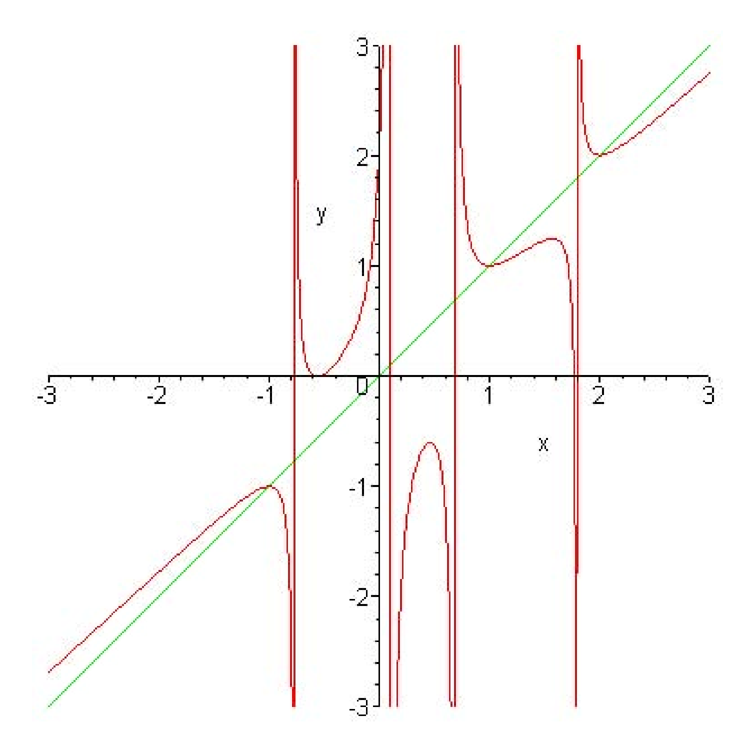


Figura 3.9: Gráfico del método de Newton para una función arbitraria.

3.5. No convergencia del método de Newton: indefinición de las iteraciones

Nuestro objetivo ahora es el de determinar la estructura topológica de $\mathcal{C}(f)$ y caracterizar la dinámica de la función de iteración N_f restringida al conjunto $\mathcal{C}(f)$. Para ello, se pueden dar los siguientes resultados, cuyas demostraciones pueden verse en [117].

Proposición 3.6. *Los conjuntos $\mathcal{B}(f)$, $\mathcal{A}(f)$ y $\mathcal{C}(f)$, definidos en (3.4), (3.6) y (3.7) respectivamente, son invariantes por la función de iteración N_f definida en (3.2).*

Mostraremos a continuación que bajo ciertas condiciones adicionales, el conjunto $\mathcal{C}(f)$ es no numerable y contiene infinitas órbitas periódicas de N_f.

Lema 3.7. *Toda banda (c_1, c_2) que contenga una raíz α de la ecuación (3.1) contiene una órbita periódica de período 2 de la función de iteración N_f definida en (3.2).*

Lema 3.8. *Sea $g : D \subset \mathbb{R} \to \mathbb{R}$. Supongamos que $I_1, \ldots, I_k$, $k \geq 2$ son intervalos compactos disjuntos dos a dos, y que para cada $j = 1, 2, \ldots, k$, la restricción de g al intervalo I_j, $g|_{I_j}$ es continua. Si para cada $m \in \mathbb{N}$ se cumple que*

$$\bigcup_{j=1}^{k} I_j \subset g(I_m),$$

entonces g tiene puntos periódicos de todos los períodos. De hecho para cada $n \in \mathbb{N}$ existen k^n puntos que satisfacen la ecuación $g^n(x) = x$, y la clausura del conjunto de los puntos periódicos es no numerable e invariante.

La siguiente proposición, es una generalización de un resultado probado por Barna en 1956 para polinomios de grado $n \geq 4$, con coeficientes reales y con n raíces reales distintas.

Proposición 3.9. *Supongamos que f satisface las condiciones **C.1** y **C.2**, y tiene al menos 4 raíces reales distintas. Entonces la función de iteración del método de Newton, N_f, definida en (3.2) tiene puntos periódicos de todos los períodos. Además, $\mathcal{A}(f) \cup \mathcal{C}(f)$ (esto es, el conjunto de puntos que no convergen a una raíz de f, equivalentemente el conjunto de puntos que no convergen a un punto fijo de N_f) es no numerable.*

Para el caso de polinomios de grado $n \geq 4$ con coeficientes reales y n raíces reales distintas, se tiene una información más precisa.

Teorema 3.10 (Barna, [11]–[14], Cosnard y Masse, [36])**.** *Sea f un polinomio de grado n con coeficientes reales. Supongamos que $n \geq 4$, y que f tiene n raíces reales distintas. Entonces*

(i) $\mathcal{C}(f)$ es un conjunto de Cantor con medida de Lebesgue cero.

(ii) $\mathcal{C}(f)$ contiene órbitas periódicas de todos los períodos de N_f. Además, para cada $x \in \mathcal{C}(f)$, $N_f'(x) < -1$.

El teorema anterior garantiza, bajo sus hipótesis, que para casi todo $x \in \mathbb{R}$ los iterados $N_f^n(x)$ convergen hacia alguna raíz de f, pues $\mathcal{C}(f) \cup \mathcal{A}(f)$ tiene medida de Lebesgue cero ($\mathcal{A}(f)$ es numerable). Más aún, el conjunto $\mathcal{C}(f)$ es repulsor, luego inaccesible computacionalmente hablando.

Definamos ahora números α y β como sigue:

(1) α es el número de bandas extremas de N_f que no contienen puntos fijos de N_f y que además son aplicadas en $\mathbb{R}$ por N_f. Es claro que $\alpha = 0, 1$ ó 2.

(2) β es el número de bandas limitadas de N_f que contienen puntos fijos de N_f.

Teorema 3.11. *Supongamos que f satisface las condiciones C.1 y C.2. Si $\alpha + \beta \geq 2$, existe un conjunto no numerable de puntos x tales que $N_f^n(x)$ no converge a ningún punto fijo de N_f. Además, si $\beta \geq 1$ entonces N_f tiene puntos periódicos de todos los períodos.*

Demostración. Si $\beta = 0$ entonces $\alpha = 2$ luego N_f no tiene puntos fijos y lo afirmado es trivial.

Si $\beta \geq 2$ el resultado se sigue del lema 3.8.

Los casos restantes son $\alpha = 1$ ó 2 y $\beta \geq 1$. En estos casos existe al menos una banda extrema B con $N_f(B) = \mathbb{R}$, y podemos elegir intervalos $I_1 \subset B$ e I_2 contenido en una banda que contiene un punto fijo de N_f, tales que $I_1 \cup I_2 \subset N_f(I_2)$ y $I_1 \cup I_2 \subset N_f(I_1)$. La prueba ahora sigue como en el lema 3.8. □

Nota 3.1. *En el caso $\alpha = 2$ y $\beta = 0$ en el teorema, hay bandas externas B_1, B_2 tales que $B_1 \cup B_2 = \emptyset$ con las propiedades $B_1 \cap B_2 \subset N_f(B_1)$ y $B_1 \cap B_2 \subset N_f(B_2)$; aplicando el lema 3.8 obtenemos el resultado.*

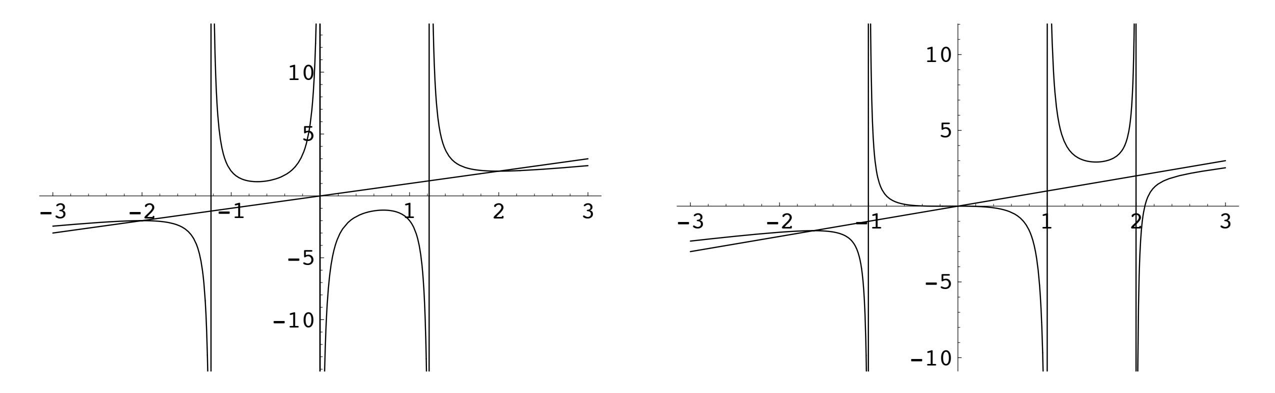


Figura 3.10: Gráficos de funciones de iteración asociadas al método de Newton. A la izquierda se muestra una situación con dos bandas extremas que contienen puntos fijos y ninguna banda interna con puntos fijos. A la derecha se muestra una situación con una banda extrema y una banda interna con puntos fijos.

3.6. No convergencia de la iteraciones del método de Newton: existencia de órbitas periódicas atractoras

Desde el punto de vista de la dinámica, los resultados anteriores son satisfactorios, pues ellos indican que para muchas funciones f, N_f definida por (3.2) tiene dinámica complicada, incluyendo una cantidad no numerable de puntos x para los cuales $N_f^k(x)$ no converge a una raíz de la ecuación $f(x) = 0$, y por otra parte este conjunto puede ser pequeño en el sentido de la medida de Lebesgue, más aún puede tener medida cero. En tal caso no debemos esperar (en sentido probabilístico) encontrar en la práctica tales puntos de no convergencia. Ahora centraremos la atención en la existencia de órbitas periódicas atractoras (recuerde que no estamos considerando como órbitas periódicas a los puntos fijos). Si N_f tiene órbitas periódicas atractoras entonces existe un intervalo abierto I tal que, si $x \in I$, entonces $N_f^k(x)$ no converge a una raíz de f. Luego los puntos de I son bien comportados desde el punto de vista de la dinámica, pero mal comportados desde el punto de vista de las iteraciones del método de Newton.

Probablemente, el resultado más antiguo en esta área es el de Barna (teorema 3.10), el cual nos dice que si f es un polinomio con coeficientes reales de grado mayor o igual que 4, con todas sus raíces reales y distintas, entonces N_f no tiene órbitas periódicas atractoras. Sin embargo, para otro tipo de funciones pueden existir órbitas periódicas atractoras. Por ejemplo, para la existencia de órbitas de período 2 basta con que se cumpla la siguiente condición para algún intervalo $[a, b]$:

1. $N_f^2([a, b]) \subset [a, b]$,

2. $|N_f'(x)| < 1$ para todo $x \in [a, b]$,

3. $|(N_f^2)'(x)| = |N_f'(N_f(x))N_f'(x)| < 1$ para todo $x \in [a, b]$.

En algunos de los ejemplos anteriores sólo mostramos los gráficos, y nos podemos plantear la siguiente pregunta. Este tipo de gráfico ¿corresponde al gráfico de N_f para alguna f?

La respuesta la obtenemos en el siguiente resultado.

Teorema 3.12. *Sea $g : \mathbb{R} \to \mathbb{R}$ que satisface*

1) g es de clase $C^2(\mathbb{R})$, excepto en un número finito de puntos $\{c_i\}_{i=1}^k$, en los cuales g no está definida, y $\lim_{x\to c_i} |g(x)| = +\infty$.

2) $\lim_{x\to c_i^+} g(x) = -\lim_{x\to c_i^-} g(x)$.

3) g tiene un número finito de puntos fijos, cada uno de los cuales es un punto crítico de g, y cualquier par de esos puntos es separado por un elemento de $\{c_i\}_{i=1}^k$.

4) g sólo tiene un número finito de puntos críticos.

Entonces existe una función $f : \mathbb{R} \to \mathbb{R}$, de clase $C^1(\mathbb{R})$, tal que $g = N_f$.

Como consecuencia del teorema anterior tenemos la siguiente proposición.

Proposición 3.13. *Dado $k > 1$, existen polinomios para los cuales la correspondientes funciones de iteración del método de Newton tienen órbitas periódicas atractoras de período k.*

Demostración. El resultado se demuestra a partir del teorema anterior, del teorema de aproximación de funciones diferenciables por polinomios (teorema de Stone-Weierstrass) y de la estabilidad de las órbitas periódicas atractoras por perturbaciones C^1. □

A continuación mostraremos una técnica que nos permite construir ciclos superatractores para el método de Newton. Aunque en este texto nos centraremos exclusivamente en el método de Newton, la técnica que vamos a desarrollar se podría extender a otros procesos iterativos.

Recordemos que una órbita periódica de longitud n o un n-ciclo para una aplicación $F : \mathbb{R} \longrightarrow \mathbb{R}$, es un conjunto finito de n puntos distintos, $\mathcal{O} = \{x_1, x_2, \ldots, x_n\}$, que satisface $F(x_i) = x_{i+1}$, para $i = 1, \ldots, n-1$ y $F(x_n) = x_1$.

Consideremos una función $f : \mathbb{R} \longrightarrow \mathbb{R}$, derivable. Al aplicar el método de Newton, obtenemos una aplicación $N_f : \mathrm{Dom}(N_f) \longrightarrow \mathbb{R}$, definida por $N_f(x) = x - f(x)/f'(x)$. Sin perdida de generalidad vamos a suponer que $N_f(x_i) = x_{i+1}$ para $i = 1, \ldots, n-1$ y $N_f(x_n) = x_1$. Si esto ocurre, claramente $\mathcal{O}$ es un n-ciclo para N_f. Ahora desde las igualdades $x_{i+1} = N_f(x_i) = x_i - f(x_i)/f'(x_i)$, para $i = 1 \ldots, n-1$ y $N_f(x_n) = x_1$, despejando nos queda

$$f'(x_i) = \frac{f(x_i)}{x_i - x_{i+1}}, \quad i = 1, \ldots, n-1, \quad \text{y} \quad f'(x_n) = \frac{f(x_n)}{x_n - x_1}. \tag{3.18}$$

Resumiendo esto, hemos probado la siguiente proposición:

Proposición 3.14. *Sea $f : \mathbb{R} \longrightarrow \mathbb{R}$ derivable y sea $\mathcal{O} = \{x_1, \ldots, x_n\} \subset Dom(N_f)$, con $x_i \neq x_j$ para todo $i, j \in \{1, \ldots, n\}$. Entonces $\mathcal{O}$ es un n-ciclo de N_f si y sólo si f satisface las ecuaciones (3.18).*

Usando técnicas de interpolación de funciones vamos a construir polinomios con la propiedad deseada. Tenemos, así, el siguiente resultado.

Proposición 3.15. *Para cada $n \geq 2$ existe un polinomio p de grado menor o igual que $2n-1$, tal que N_p tiene un n-ciclo.*

Demostración. Consideremos n puntos $x_1, x_2, \ldots, x_n$ con la propiedad que $x_i \neq x_j$ si $i \neq j$ para todo $i, j \in \{1, 2, \ldots, n\}$ y sean $y_1, y_2, \ldots, y_n$ otros n puntos con la propiedad que $y_k \neq 0$ para todo $k = 1, 2, \ldots, n$ (nótese que los puntos y_j no tienen porqué ser necesariamente distintos). Por la proposición 3.14, si $f : \mathbb{R} \longrightarrow \mathbb{R}$ satisface $f(x_i) = y_i$ para $i = 1, \ldots, n$ y las condiciones (3.18), entonces el conjunto $\mathcal{O} = \{x_1, x_2, \ldots, x_n\}$ es un n-ciclo para N_f.

Para construir el polinomio p usaremos interpolación de Hermite, que sirve para determinar un polinomio que coincide con una función dada y con su derivada primera en una serie de puntos prefijados. En concreto, buscamos un polinomio p de grado a lo más $2n-1$ y que satisface las condiciones

$$\begin{cases} f(x_i) = y_i, \ i = 1, \ldots, n, \\ \\ f'(x_i) = \dfrac{y_i}{x_i - x_{i+1}}, \ i = 1, \ldots, n-1, \ \text{y} \\ \\ f'(x_n) = \dfrac{y_n}{x_n - x_1}. \end{cases} \tag{3.19}$$

Aunque existen formas de escribir el polinomio de interpolación de Hermite en términos de las diferencias divididas de la función a interpolar o de los correspondientes polinomios de Lagrange (véase [50], por ejemplo), en esta sección emplearemos una formulación alternativa para representar dicho polinomio interpolador. Así, escribimos p como

$$p(x) = a_1 p_1(x) + a_2 p_2(x) + \cdots + a_{2n} p_{2n}(x) \tag{3.20}$$

donde cada p_i es un polinomio de grado $i-1$, para $i = 1, \dots, 2n$, y son definidos por

$$\begin{cases} p_1(x) &= 1 \\ p_2(x) &= p_1(x)(x-x_1) = (x-x_1) \\ p_3(x) &= p_2(x)(x-x_1) = (x-x_1)^2 \\ p_4(x) &= p_3(x)(x-x_2) = (x-x_1)^2(x-x_2) \\ p_5(x) &= p_4(x)(x-x_2) = (x-x_1)^2(x-x_2)^2 \\ &\vdots \\ p_{2i-1}(x) &= p_{2i-2}(x-x_{i-1}) \\ p_{2i}(x) &= p_{2i-1}(x)(x-x_i) \\ &\vdots \\ p_{2n-1}(x) &= p_{2n-2}(x)(x-x_{n-1}) \\ p_{2n}(x) &= p_{2n-1}(x)(x-x_n). \end{cases} \tag{3.21}$$

Notemos que $p_{2i-1}(x_i) = p_{2i-2}(x_i)(x_i - x_{i-1}) \neq 0$ y que $p_{2i}(x_i) = p_{2i-1}(x_i)(x_i - x_i) = 0$, esto nos dice que $p_j(x_i) \neq 0$ para $j \leq 2i-1$ y que $p_j(x_i) = 0$ para $j \geq 2i$.

Por otra parte, tenemos que $p'_{2i}(x) = p'_{2i-1}(x)(x-x_i) + p_{2i-1}(x)$, de donde $p'_{2i}(x_i) = p_{2i-1}(x_i) \neq 0$ y $p'_{2i+1}(x) = p'_{2i}(x)(x-x_i) + p_{2i}(x)$, así $p'_{2i+1}(x_i) = p_{2i}(x_i) = 0$. Luego, $p'_j(x_i) \neq 0$ para $j \leq 2i$ y $p'_j(x_i) = 0$ para $j \geq 2i+1$. Para determinar el polinomio $p(x)$ debemos ser capaces de calcular los coeficientes a_i, para $i = 1, \dots, 2n$. Para ello, debemos resolver un sistema de $2n$ ecuaciones con $2n$ incógnitas, si es posible hacerlo. El sistema de ecuaciones asociado a nuestro problema es un sistema dado por una matriz A, triangular inferior, cuyas filas A_j, para $j = 1, \dots, 2n$ están definidas de la siguiente manera.

$$\begin{aligned} A_{2i-1} &= (p_1(x_i) \;\; p_2(x_i) \;\; p_3(x_i) \;\; \cdots \;\; p_{2i-1}(x_i) \;\; 0 \;\; 0 \;\; 0 \;\; \cdots \;\; 0) \\ A_{2i} &= (p'_1(x_i) \;\; p'_2(x_i) \;\; p'_3(x_i) \;\; \cdots \;\; p'_{2i-1}(x_i) \;\; p'_{2i}(x_i) \;\; 0 \;\; 0 \;\; \cdots \;\; 0) \end{aligned}$$

con $i = 1, \dots, n$.

Así el sistema de ecuaciones lineales a resolver viene dado por $A\mathbf{X} = b$, donde

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_{2n} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} a_1 \\ \vdots \\ a_{2n} \end{pmatrix} \quad \text{y} \quad b = \begin{pmatrix} y_1 \\ \frac{y_1}{x_1 - x_2} \\ \vdots \\ y_n \\ \frac{y_n}{x_n - x_1} \end{pmatrix}. \tag{3.22}$$

Este sistema tiene solución, pues el determinante de la matriz A es no nulo. Para ello basta notar que la matriz A es triangular inferior y que las componente de su diagonal son los elementos de la forma $p_{2i-1}(x_i)$ y $p'_{2i}(x_i)$, los cuales son no nulos para $i = 1, \dots, n$. Esto completa la prueba de la proposición. □

Ejemplo 3.7. *En este ejemplo vamos a construir, usando la técnica desarrollada en la proposición 3.15, un polinomio cuyo método de Newton asociado tenga una órbita periódica de período 3.*

Para ello consideremos los datos

$$\begin{cases} x_1 = 0, & y_1 = 1, \\ x_2 = 1, & y_2 = -1, \\ x_3 = 2, & y_3 = 1. \end{cases}$$

Es decir, estamos imponiendo al polinomio $p(x)$ que $p(0) = 1$, $p(1) = -1$ y $p(2) = 1$ y que sus derivadas en esos puntos satisfagan

$$p'(0) = \frac{1}{0-1} = -1, \quad p'(1) = \frac{-1}{1-2} = 1, \quad p'(2) = \frac{1}{2-0} = \frac{1}{2}.$$

Ahora debemos construir los polinomios $p_i(x)$, $i = 1, \ldots, 6$, los cuales son dados como sigue:

$$\begin{cases} p_1(x) &=& 1 \\ p_2(x) &=& p_1(x)(x - x_1) = x \\ p_3(x) &=& p_2(x)(x - x_1) = x^2 \\ p_4(x) &=& p_3(x)(x - x_2) = x^2(x-1) \\ p_5(x) &=& p_4(x)(x - x_2) = x^2(x-1)^2 \\ p_6(x) &=& p_5(x)(x - x_3) = x^2(x-1)^2(x-2). \end{cases}$$

El polinomio buscado tiene la forma

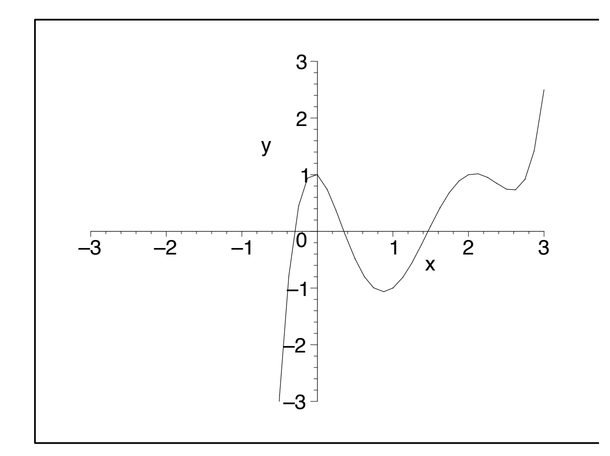
$$p(x) = a_1 + a_2 x + a_3 x^2 + a_4 x^2(x-1) + a_5 x^2(x-1)^2 + a_6 x^2(x-1)^2(x-2).$$

En este caso, las filas de la matriz A del sistema de ecuaciones lineales $A\mathbf{X} = b$ son:

$$\begin{array}{lcllllll} A_1 &=& p_1(x_1) & 0 & 0 & 0 & 0 & 0 \\ A_2 &=& p_1'(x_1) & p_2'(x_1) & 0 & 0 & 0 & 0 \\ A_3 &=& p_1(x_2) & p_2(x_2) & p_3(x_2) & 0 & 0 & 0 \\ A_4 &=& p_1'(x_2) & P_2'(x_2) & p_3'(x_2) & p_4'(x_2) & 0 & 0 \\ A_5 &=& p_1(x_3) & p_2(x_3) & p_3(x_3) & p_4(x_3) & p_5(x_3) & 0 \\ A_6 &=& p_1(x_3)' & p_2'(x_3) & p_3'(x_3) & p_4'(x_3) & p_5'(x_3) & p_6'(x_3). \end{array}$$

Calculando las derivadas de los polinomios $p_i(x)$, $i = 1, \ldots, 6$ y evaluando $p_i(x)$y $p_i'(x)$, obtenemos el sistema de ecuaciones lineales:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 2 & 4 & 4 & 4 & 0 \\ 0 & 1 & 4 & 8 & 12 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ \frac{1}{2} \end{pmatrix}$$



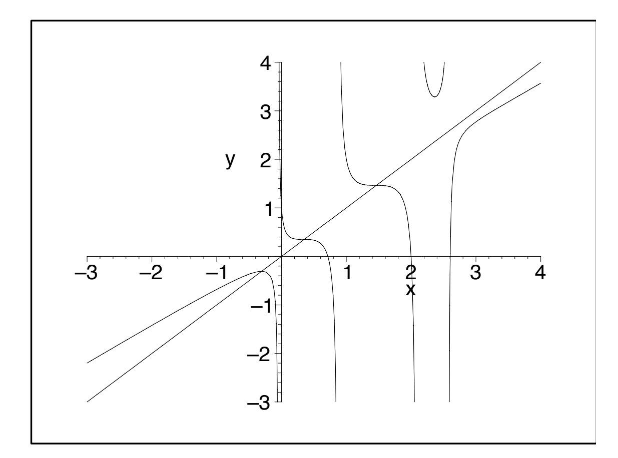
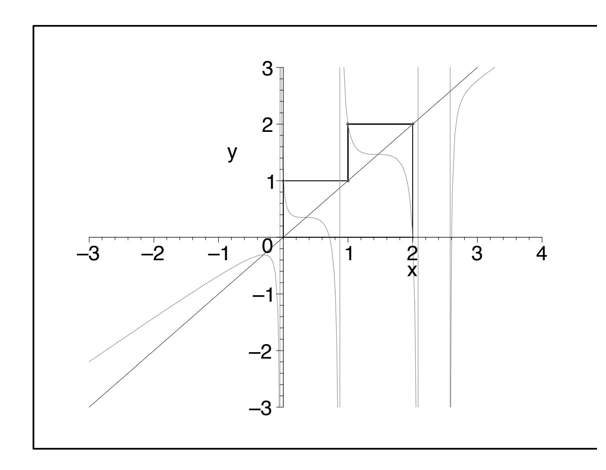


Figura 3.11: Gráficas del polinomio $p(x)$ definido en (3.23) y de su correspondiente función de iteración para el método de Newton $N_p(x)$.



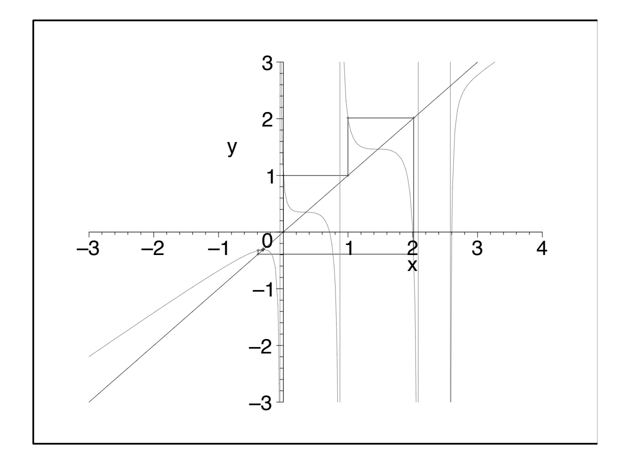


Figura 3.12: Órbitas de los puntos $x_0 = 0$ y $x_0 = 0.0001$ para el método de Newton asociado al polinomio $p(x)$ definido en (3.23).

cuya solución es $a_1 = 1$, $a_2 = -1$, $a_3 = -1$, $a_4 = 4$, $a_5 = -5/2$, y $a_6 = 7/8$. Por lo tanto, el polinomio buscado es

$$p(x) = 1 - x - x^2 + 4x^2(x-1) - \frac{5}{2}x^2(x-1)^2 + \frac{7}{8}x^2(x-1)^2(x-2). \quad (3.23)$$

En las figuras 3.11 y 3.12 se muestran los gráficos del polinomio p construido, del método de Newton asociado, N_p, con el 3-ciclo $\mathcal{O} = \{0, 1, 2\}$. También se muestra la órbita del punto $x_0 = 0.0001$, indicando el carácter repulsor del 3-ciclo $\{0, 1, 2\}$.

Ahora veremos como obtener a partir del polinomio $p(x)$ construido en la proposición 3.15 otro polinomio tal que el n-ciclo dado por la proposción sea superatractor. Pero antes de eso veamos como caracterizamos el hecho que un n-ciclo sea superatractor para el método de Newton.

Proposición 3.16. *Sea $p(x)$ un polinomio cuyo método de Newton tiene un n-ciclo, $\mathcal{O} = \{x_1, x_2, \ldots, x_n\}$. Si $p''(x_i) = 0$ para algún $i \in \{1, 2, \ldots, n\}$, entonces $\mathcal{O}$ es superatractor.*

Demostración. Sin pérdida de generalidad, podemos suponer que $p''(x_1) = 0$. Como $N_p(x_i) = x_{i+1}$ para $i = 1, \ldots, n-1$ y $N_p(x_n) = x_1$, aplicando la regla de la cadena se tiene

$$(N_p^k)'(x) = N_p'(N_p^{k-1}(x))N_p'(N_p^{k-2}(x)) \cdots N_p'(N_p(x))N_p'(x).$$

Tenemos entonces

$$(N_p^n)'(x_1) = N_p'(x_n)N_p'(x_{n-1}) \cdots N_p'(x_2)N_p'(x_1).$$

Como $N_p'(x) = p(x)p''(x)/p'(x)^2$ y estamos suponiendo que $p''(x_1) = 0$, en la expresión para $(N_p^n)'(x_1)$ el factor $N_p'(x_1)$ se anula, por lo tanto $(N_p^n)'(x_1) = 0$, y el resultado está probado.

Ahora tenemos el siguiente resultado que nos permite construir ciclos atractores para el método de Newton. □

Teorema 3.17. *Para cada $n \geq 2$ existe un polinomio $p(x)$ de grado menor o igual a $2n$ de modo que $N_p(x)$ tiene un n-ciclo superatractor.*

Demostración. Dado un conjunto $\mathcal{O} = \{x_1, \ldots, x_n\}$ de n puntos distintos entre sí, y dado otro conjunto de n puntos $\{y_1, \ldots, y_n\}$, usando la proposición 3.15 construimos un polinomio $p(x)$ tal que $\mathcal{O}$ es un n-ciclo para $N_p(x)$. El polinomio $p(x)$ tiene la forma $p(x) = a_1p_1(x) + \cdots + a_{2n}p_{2n}(x)$. Definimos un nuevo polinomio

$$\tilde{p}(x) = p(x) + a_{2n+1}p_{2n+1}(x),$$

donde a_{2n+1} es un paramétro a determinar y $p_{2n+1}(x) = p_{2n}(x)(x - x_n)$. Para la determinación de a_{2n+1} imponemos la condición de que el n-ciclo $\mathcal{O}$ sea superatractor para $N_{\tilde{p}}(x)$. Notemos primero que la nueva condición no altera $\mathcal{O}$, ya que tenemos $p_{2n+1}(x_i) = 0$ para $i = 1, \ldots, n$. Como queremos que $\mathcal{O}$ sea superatractora para $N_{\tilde{p}}(x)$, por la proposición 3.16, nos basta encontrar a_{2n+1} de modo que $p''(x_1) = 0$, imponiendo esta condición y resolviendo para a_{2n+1} se tiene lo pedido. □

Ejemplo 3.8. *Como continuación del ejemplo 3.7 vamos a construir ahora un polinomio cuyo método de Newton asociado tenga una órbita periódica de período 3 que además sea superatractora.*

Partimos del polinomio $p(x)$ definido en (3.23) y definimos

$$\tilde{p}(x) = 1 - x - x^2 + 4x^2(x-1) - \frac{5}{2}\,x^2(x-1)^2 + \frac{7}{8}\,x^2(x-1)^2(x-2) + a_7\,x^2(x-1)^2(x-2)^2.$$

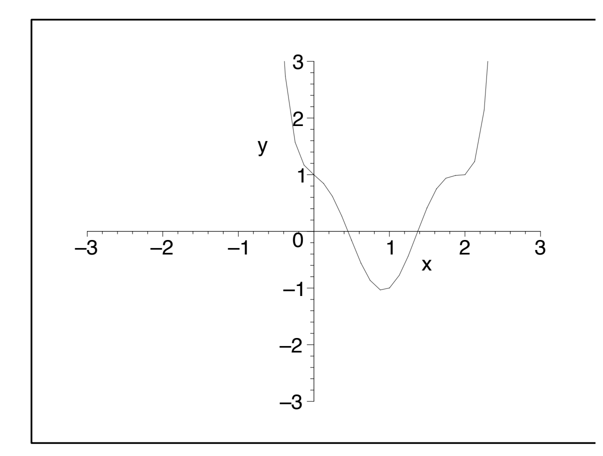
Haciendo $\tilde{p}\,''(0) = 0$, obtenemos $a_7 = 37/16$. Luego el polinomio es de la forma

$$\tilde{p}(x) = 1 - x - \frac{115}{8}x^3 + \frac{385}{16}x^4 - 13x^5 + \frac{37}{16}x^6 \tag{3.24}$$

es tal que

$$N_{\tilde{p}} = \frac{-460x^3 + 1155x^4 - 832x^5 + 185x^6 - 16}{2(-8 - 345x^2 + 770x^3 - 520x^4 + 111x^5)} \tag{3.25}$$

tiene a $\mathcal{O} = \{0, 1, 2\}$ como un 3-ciclo superatractor.



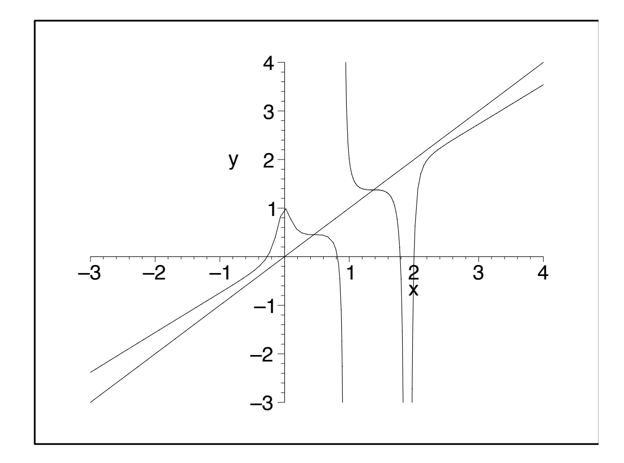
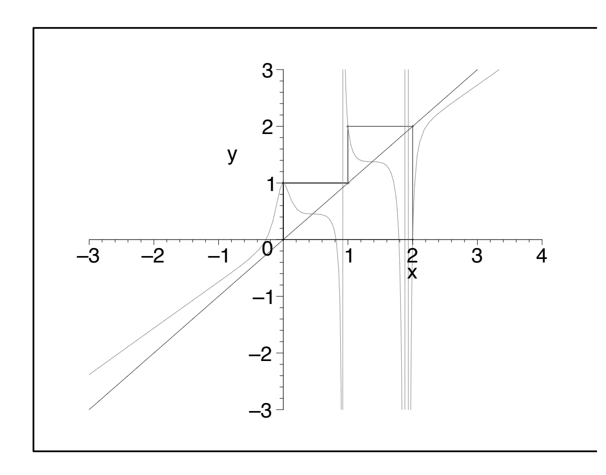


Figura 3.13: Gráficas del polinomio $\tilde{p}(x)$ definido en (3.24) y de su correspondiente función de iteración para el método de Newton $N_{\tilde{p}}(x)$, definida en (3.25).



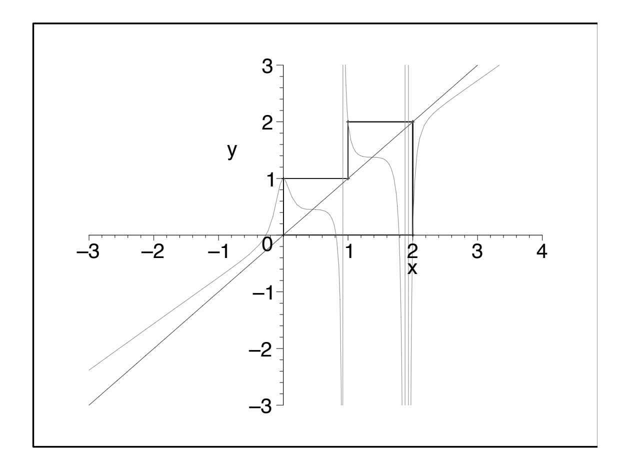


Figura 3.14: Órbitas de los puntos $x_0 = 0$ y $x_0 = 0.00003$ para el método de Newton asociado al polinomio $\tilde{p}(x)$ definido en (3.24).

3.7. Bifurcaciones en el método de Newton

En lo que sigue estudiaremos la aparición de bifurcaciones en la función de iteración del método de Newton (3.2). Esto ocurre cuando la función f pierde una raíz pasando a través de una tangencia cuadrática con el eje OX.

En lo que sigue, supondremos que $g : \mathbb{R} \to \mathbb{R}$ es una función dos veces derivable en un entorno del origen, que cumple las siguientes condiciones:

$$g(0) = 0, \qquad g'(0) = 0, \qquad g''(0) > 0. \tag{3.26}$$

Para $\mu \in \mathbb{R}$, definimos ahora $f_\mu(x) = g(x) + \mu$. Localmente, en una vecindad de 0, las funciones f_μ tienen unos gráficos como los que se muestran en la figura 3.15.

Denotemos a su función de iteración del método de Newton N_{f_μ} simplemente por N_μ. Observemos que f_μ es una traslación de μ unidades en la dirección del eje OY de la función g. Localmente, las gráficas de N_μ tienen la apariencia que se muestra en la figura 3.16.

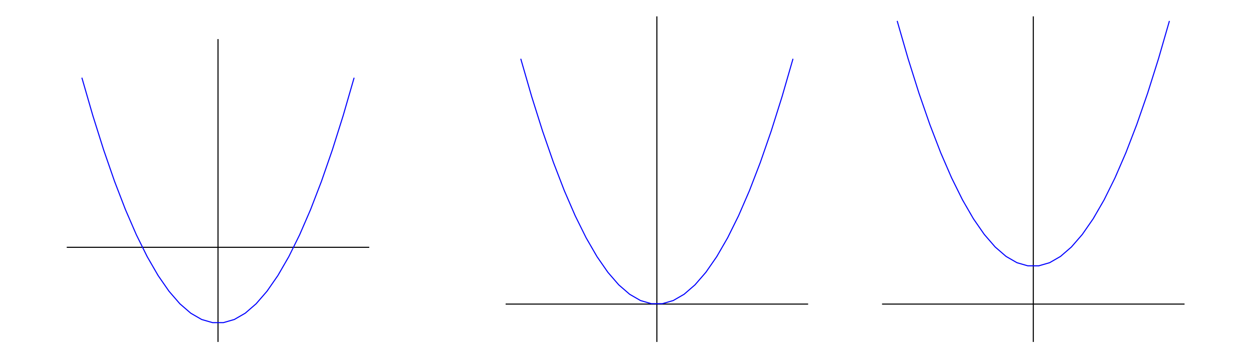


Figura 3.15: Gráficos de las funciones $f_\mu(x) = g(x) + \mu$ para $\mu < 0$, $\mu = 0$ y $\mu > 0$.

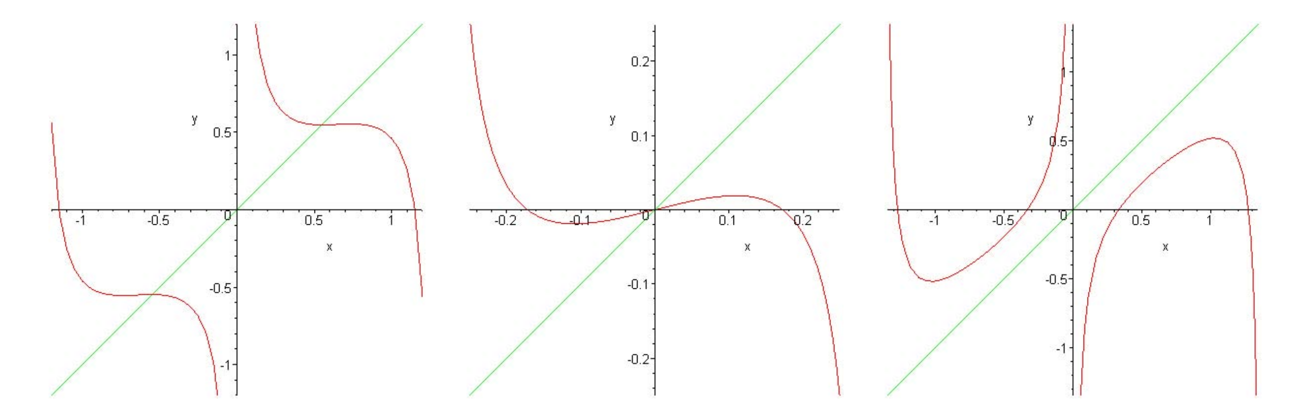


Figura 3.16: Gráfica de la función N_μ para $\mu < 0$, $\mu = 0$ y $\mu > 0$.

Obsérvese que $f'_\mu(x) = g'(x)$. Por tanto,

$$N_\mu(x) = x - \frac{g(x)+\mu}{g'(x)} = N_g(x) - \frac{\mu}{g'(x)} \tag{3.27}$$

y además

$$N'_\mu(x) = N'_g(x) + \frac{\mu g''(x)}{(g'(x))^2} = \frac{(g(x)+\mu)g''(x)}{(g'(x))^2}. \tag{3.28}$$

Si $g'(x) \neq 0$ entonces

$$\frac{\partial}{\partial \mu} N_\mu(x) = -\frac{1}{g'(x)}.$$

Nótese que esta derivada no depende del parámetro μ, y está bien definida cuando $g'(x) \neq 0$. Por otra parte, existen puntos excepcionales (x_0, μ_0) donde $\frac{\partial}{\partial \mu} N_\mu(x)$ no existe, por ejemplo, en $(x_0, \mu_0) = (0, \mu)$. Esos puntos corresponden a puntos donde $g'(x_0) = 0$. Note que las bandas de N_μ son, esencialmente, independiente de μ y son las mismas que para N_g; las únicas excepciones ocurren en los valores del parámetro μ en los cuales f_μ tiene una raíz que es también un punto crítico, es decir, para (x_0, μ_0) tales que $f_{\mu_0}(x_0) = 0$ y $f'_{\mu_0}(x_0) = 0$, es decir, $0 = f_{\mu_0}(x_0) = g(x_0) + \mu_0$ y $f'_{\mu_0}(x_0) = g'(x_0) = 0$, o más específicamente

$$\begin{cases} g'(x_0) = 0, \\ \mu_0 = -g(x_0). \end{cases}$$

Después de estas consideraciones, podemos dar el siguiente lema.

Lema 3.18. *Sea I un intervalo compacto sobre el cual g' no se anula, entonces cuando $\mu \to 0$ se tiene que N_μ converge a N_g y N'_μ converge a N'_g, siendo la convergencia en ambos casos uniforme.*

Demostración. La prueba se sigue trivialmente de las igualdades (3.27) y (3.28). □

Para encontrar puntos periódicos atractores de una aplicación 1-dimensional (real o compleja) es conveniente seguir la órbitas de los puntos críticos de la aplicación. Para nuestro caso, existen al menos dos razones para esto: primero si un punto crítico de N_μ es periódico de período k, entonces su órbita es superatractora; segundo si N_μ es una función racional (cociente de dos polinomios sin factores comunes), cosa que ocurre para N_g por ejemplo cuando g es un polinomio, entonces existe una fuerte conexión entre las órbitas periódicas atractoras y los puntos críticos de la función de iteración del método de Newton. De hecho, en general, el teorema de Julia ([16, p. 194]) establece que cada órbita periódica atractora de una función racional arbitraria contiene un punto crítico en su cuenca de atracción inmediata.

En lo que sigue estudiaremos las órbitas de los puntos críticos de N_μ cuando el gráfico de f_μ pasa a través de una tangencia cuadrática con el eje OX en el punto $(x, \mu) = (0, 0)$. Asumiremos que tal punto crítico existe. Tenemos que

$$N'_\mu(x) = \frac{f_\mu(x) f''_\mu(x)}{(f'_\mu(x))^2},$$

luego los puntos críticos de N_μ, caso $f'_\mu(x) \neq 0$, son los puntos donde $f_\mu(x) = 0$ (raíces de f_μ) o $f''_\mu(x) = 0$ (puntos de inflexión de f). Cuando un punto crítico de N_μ no es una raíz de $f_\mu(x) = 0$ diremos que es *punto crítico libre*. Como $f''_\mu = g''$ se tiene que los puntos críticos libres de N_μ son aquellos donde g'' se anula, es decir, genéricamente, un punto de inflexión de g. Como $g''(x) > 0$ en una vecindad de $x = 0$, la condición de existencia de un punto crítico libre se puede formular de la siguiente forma: existen puntos $p < 0 < q$ tales que $g''(p) = 0$ y $g''(x) > 0$ sobre (p, q). Esto junto con la condición que impusimos a g, es decir, $g(0) = 0$, $g'(0) = 0$, $g''(0) > 0$, y el teorema del valor medio implican que el único punto sobre (p, q) donde o bien g o g' se anulan es $x = 0$, $g(x) \geq 0$ sobre (p, q), $g'(x) < 0$ sobre $(p, 0)$ y $g'(x) > 0$ sobre $(0, q)$. Ahora, como $N'_\mu(x) = N'_g(x) + \dfrac{\mu g''(x)}{g'(x)^2}$ se tiene que $N'_\mu(x) > N'_g(x) > 0$ sobre $(p, q) - \{0\}$ y $\mu > 0$. Por otra parte, dado que $\dfrac{\partial}{\partial \mu} N_\mu(x) = -\dfrac{1}{g'(x)}$ se sigue que $\dfrac{\partial}{\partial \mu} N_\mu(x) > 0$ sobre $(p, 0)$ y $\dfrac{\partial}{\partial \mu} N_\mu(x) < 0$ sobre $(0, q)$. Tenemos el siguiente lema.

Lema 3.19. *Supongamos que $\mu > 0$. Si $N^i_\mu(p) < 0$ para $0 \leq i \leq k$ entonces*

$$\frac{\partial}{\partial \mu} N^k_\mu(p) > 0.$$

Demostración. Por la regla de la cadena, tenemos que

$$\frac{\partial}{\partial\mu}N_\mu^j(x) = \prod_{\ell=0}^{j-1}\frac{\partial}{\partial\mu}N_\mu(N_\mu^\ell(x)).$$

Ahora como $N_\mu^i(p) < 0$ para $0 \leq i \leq k$ se tiene que

$$\frac{\partial}{\partial\mu}N_\mu(N_\mu(p)) > 0.$$

Por lo tanto,

$$\frac{\partial}{\partial\mu}N_\mu^k(p) = \prod_{\ell=0}^{k-1}\frac{\partial}{\partial\mu}N_\mu(N_\mu^\ell(p)) > 0,$$

como deseábamos probar. □

Las hipótesis del lema anterior se satisfacen para valores pequeños de μ. Además por la continuidad respecto a μ podemos enunciar el siguiente resultado.

Lema 3.20. *Dado $k > 0$ existe $\mu_0 > 0$ tal que si $|\mu| > \mu_0$ entonces $p < N_\mu(p) < N_\mu^2(p) < \cdots < N_\mu^k(p) < 0$.*

Demostración. Como $N_g'(x) > 0$ sobre $(p,q)-\{0\}$ ($N_g'(0) = 1/2$) se sigue que N_g es monótona creciente sobre $[p,q]$, y para cualquier $x \in [p,q]$, $N_g(x)$ está entre x y 0 luego, para cualquier $x \in [p,q]$ la sucesión $N_g^n(x)$ converge monótonamente a 0. Además, para cualquier constante positiva β, se tiene que N_μ converge C^1 uniformemente a N_g sobre $[p,-\beta] \cup [\beta, q]$ cuando $\mu \to 0$, así si elegimos $\beta < \text{mín}\{N_g^k(q), |N_g^k(p)|\}$ encontramos μ_0 como asegura el lema. □

Lema 3.21. *Dado $k > 0$ existe $\mu_1 > 0$ tal que si $0 < \mu < \mu_1$ entonces*

(1) $p < N_\mu(p) < \cdots < N_\mu^{k-1}(p) < 0$;

(2) $N_\mu^k(p) > N_\mu^{k-1}(p)$;

(3) $N_\mu^{k-1}(p) \to 0$ cuando $\mu \nearrow \mu_1$, luego, $N_\mu^k(p) \to \infty$ cuando $\mu \nearrow \mu_1$.

Demostración. Es una consecuencia inmediata de las condiciones y de los lemas anteriores. La idea es que si la condición (1) vale, el conjunto $\{N_\mu^i(p)\}$ es monótono en i para $0 \leq i \leq k$ y en μ para $\mu > 0$.

Por los dos últimos lemas, para cada $k \geq 1$ existe un μ-intervalo $[a_k, b_k]$ tal que si $\mu \in [a_k, b_k]$ entonces

$$\begin{cases} N_\mu^i(p) \in (p,0), & 1 \leq i \leq k, \\ N_\mu^k(p) \in [0,q], & \\ N_{a_k}^k(p) = 0, & \\ N_{b_k}^k(p) = q. & \end{cases}$$

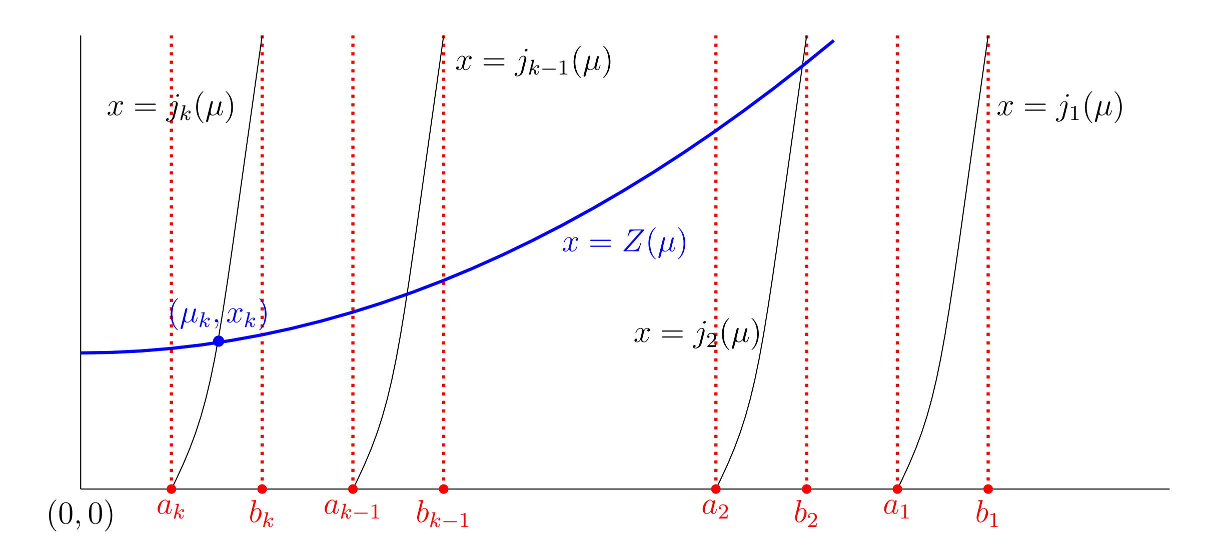


Figura 3.17: Gráficas de las curvas Z y j_k que aparecen en la demostración del lema 3.21.

Además, para cada k se tiene que $a_k > b_{k+1} > 0$ y $\lim_{k\to\infty} a_k = 0$. Definamos la aplicación $j_k : [a_k, b_k] \to [0, q]$ por $j_k(\mu) = N_\mu^k(p)$. Cada aplicación j_k es monótona creciente y aplica $[a_k, b_k]$ sobre $[0, q]$.

Para valores pequeños y positivos de μ, se tiene que N_μ aplica el intervalo $[0, q]$ sobre el intervalo $[p, 0]$; luego el conjunto $N_\mu^{-1}(p) \cap (0, q)$ es no vacío. Por el teorema de la función implícita se tiene que el conjunto $N_\mu^{-1}(p) \cap (0, q)$ es el gráfico de una función creciente $x = Z(\mu)$. Las curvas Z y j_k se muestran en la figura 3.17.

Para todo k grande existe un punto (μ_k, x_k) satisfaciendo

$$\begin{cases} x_k = Z(\mu_k) \\ x_k = j_k(\mu_k). \end{cases}$$

Consecuentemente, para $\mu = \mu_k$ el punto crítico p de N_μ está sobre una órbita periódica superatractora de período $k + 1$ para N_μ. Tenemos $x_k = Z(\mu_k) = N_{\mu_k}^{-1}(p) \cap (0, q)$, de donde $N_{\mu_k}(x_k) = \{p\} \cap N_{\mu_k}((0, q))$ y $x_k = j_k(\mu_k)$. Luego, $p = N_{\mu_k}(x_k) = N_{\mu_k}^{k+1}(p)$, y como $g''(p) = 0$ se tiene que $(N_{\mu_k}^{k+1})'(p) = 0$. □

Tenemos el siguiente teorema.

Teorema 3.22. *Supongamos que g es de clase al menos $C^3(I)$ donde I es un intervalo que contiene al 0 y además se satisfacen las siguientes condiciones: $g(0) = 0,\ g'(0) = 0,\ g''(0) \neq 0$.*

Sea $f_\mu(x) = g(x) + \mu$. Si existe un punto p con $g''(p) = 0$ entonces para todo k suficientemente grande, existe un intervalo abierto I_k sobre el μ–eje tal que

(1) Para cada $\mu \in I_k$, la aplicación de Newton N_μ tiene una órbita periódica atractora de período k;

(2) *cada intervalo I_k contiene al menos un punto μ_k tal que p es un punto periódico super-atractor de N_μ para $\mu = \mu_k$;*

(3) *si $I_k = (L_k, R_k)$ entonces L_k y R_k tienden a 0 cuando $k \to \infty$.*

Demostración. En lo anterior hemos tratado el caso $p < 0$ y $g''(p) > 0$. En esa condiciones hemos demostrado la existencia de una órbita periódica superatractora para $\mu = \mu_k$. El teorema de la función implícita muestra que esa órbita continua existiendo para todos los valores de μ suficientemente cercanos a μ_k. Sea I_k un intervalo que contiene a μ_k y tal que para cada $\mu \in I_k$, N_μ tiene una órbita periódica atractora de período k que contiene a p en su cuenca de atracción. Si $p > 0$ el argumento es esencialmente el mismo. Para el caso $g''(0) < 0$, notemos que la aplicación de Newton N_μ de $g + \mu$ es exactamente la misma que para $-g - \mu$ la cual le podemos aplicar el argumento anterior. Esto completa la prueba del teorema. □

Para ilustrar lo discutido anteriormente el lector puede considerar el ejemplo siguiente. Sea $f(x) = x^3 - 3x + \mu$. Esta familia de aplicaciones tiene una tangencia cuadrática con el eje x en $x = 1$ cuando $\mu = 2$. Tenemos $f'_\mu(x) = 3x^2 - 3$, $f''_\mu(x) = 6x$. Luego el único cero de f''_μ es $x = 0$. Por lo tanto, del teorema de Julia N_μ si tiene una órbita periódica de período mayor que 1, entonces los iterados de 0 deben tender a esa órbita. Un pequeño experimento (véase el cuadro 3.1), nos permite calcular los valores del parámetro μ_k para los cuales la función de iteración N_μ definida en (3.27) tiene un ciclo de período k.

Cuadro 3.1: Ciclos periódicos y valores del parámetro μ_k para la función N_μ definida en (3.27).

Período k	Parámetro μ_k
2	3.674234614
3	7.7513408204
4	15.556213393
5	30.155532349

Capítulo 4

Dinámica del método de Newton en el plano complejo

4.1. Antecedentes: el problema de Cayley

El estudio de la dinámica del método de Newton en el campo complejo tiene una importancia histórica, desde que E. Schröder (1870) y A. Cayley (1879) propusieran usar el método de Newton para resolver ecuaciones definidas en el plano complejo:

$$f(z) = 0, \quad f : \mathbb{C} \to \mathbb{C}. \tag{4.1}$$

El que se conoce como *problema de Cayley* consiste en estudiar las cuencas de atracción del método de Newton cuando es aplicado para aproximar las raíces del polinomio complejo $p(z)$. En palabras del propio Cayley, el problema se puede formular como sigue:

> «...the problem is to determine the regions of the plane, such that P [initial point] being taken at pleasure anywhere within one region we arrive ultimately at the point A [a root of the polynomial]...»

Con la notación actual, si z^* es una solución de (4.1), y

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} \tag{4.2}$$

es la sucesión generada por el método de Newton a partir de un cierto $z_0 \in \mathbb{C}$, se trata de caracterizar la región

$$A(z^*) = \{z_0 \in \mathbb{C} : z_n \to z^*\},$$

conocida como *cuenca de atracción* de la raíz z^*. Cayley ([28], [29]) consiguió caracterizar los cuencas de atracción de las raíces de un polinomio cuadrático, en concreto del polinomio $p(z) = z^2 - 1$, aunque fracasó en su intento de extender el estudio al caso cúbico y a grados superiores. Después de un tiempo tratando de resolver el problema para el polinomio cúbico $p(z) = z^3 - 1$, escribe en [30] la siguiente sentencia:

> «J'espère appliquer cette théorie au cas d'une équation cubique, mais les calculs sont beaucoup plus difficiles»

En la figura 4.1 mostramos las cuencas de atracción de los polinomios $p(z) = z^2 - 1$ y $p(z) = z^3 - 1$. En el primer caso vemos que los puntos de partida situados en el semiplano $\mathbb{C}^- = \{z \in \mathbb{C} : \text{Re}(z) < 0\}$ convergen a la raíz $z^* = -1$, mientras que los puntos de partida situados en el semiplano $\mathbb{C}^+ = \{z \in \mathbb{C} : \text{Re}(z) > 0\}$ convergen a la raíz $z^* = 1$. En la separación de ambas regiones, el eje imaginario, en donde el método de Newton presenta un comportamiento caótico. Sin embargo, como vemos en la segunda gráfica de la figura 4.1, la situación para el polinomio $p(z) = z^3 - 1$ es mucho más complicada. La separación entre las cuencas de atracción de las tres raíces, 1, $(-1 + \sqrt{3}i)/2$ y $(-1 - \sqrt{3}i)/2$, no es tan diáfana y tiene una estructura mucho más enrevesada.

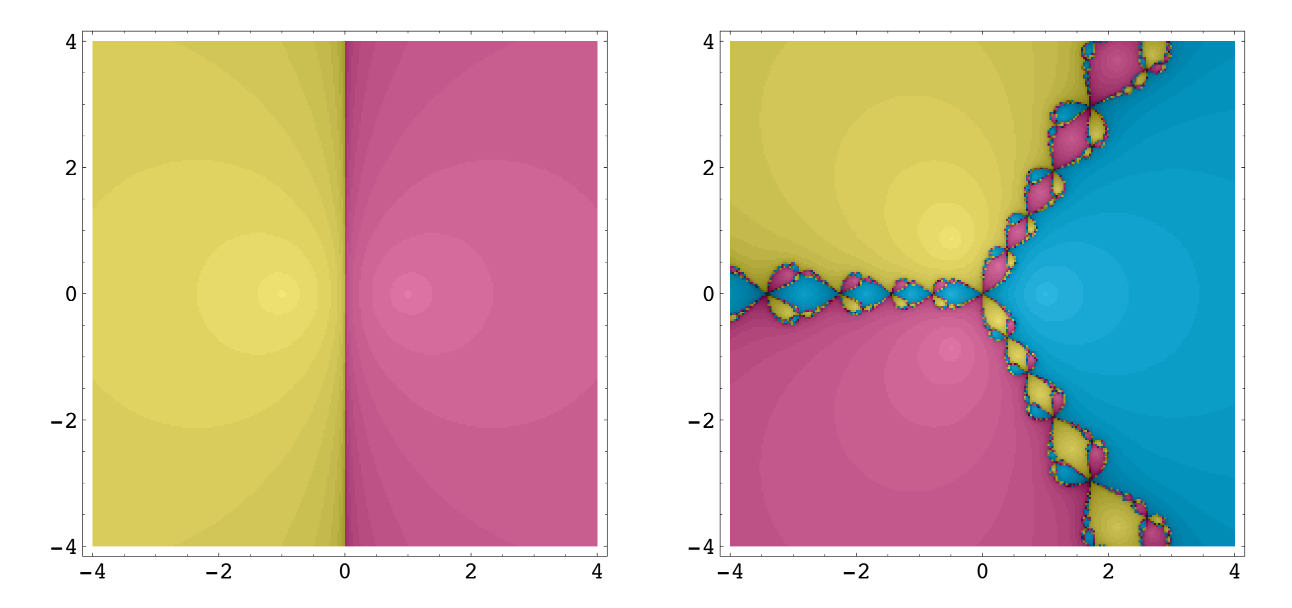


Figura 4.1: A la izquierda se muestran las cuencas de atracción del polinomio $p(z) = z^2 - 1$ y a la derecha las del polinomio $p(z) = z^3 - 1$. Las regiones pintadas con el mismo color están formadas por puntos de partida para los cuales el método de Newton converge a la misma raíz del polinomio correspondiente.

No es de extrañar, por tanto, que Cayley, que no disponía de los potentes programas de dibujo y cálculo simbólico que tenemos en la actualidad, encontrara dificultades al intentar pasar de una ecuación cuadrática a una cúbica.

Posteriormente, G. Julia y P. Fatou consideran funciones racionales en una forma más general, obteniendo resultados significativos y sientan el estudio de iteraciones de funciones racionales en el plano complejo extendido, también conocido como la esfera de Riemann. En otras palabras, con sus trabajos comienza en forma sistemática el estudio de los sistemas dinámicos complejos. La base de sus trabajos fue el estudio realizado por Montel sobre familias normales.

4.2. Conceptos básicos de dinámica compleja

En esta sección revisamos algunos conceptos básicos de dinámica compleja que serán necesarios para el análisis del método de Newton cuando se aplica a funciones definidas en el plano complejo. En especial, haremos hincapié en el caso de la aplicación del método de Newton a ecuaciones polinómicas.

En primer lugar, notemos que dado un polinomio definido en el plano complejo $p : \mathbb{C} \to \mathbb{C}$, la función de iteración del método de Newton

$$N_p(z) = z - \frac{p(z)}{p'(z)} = \frac{zp'(z) - p(z)}{p'(z)} \tag{4.3}$$

puede considerarse como una función racional definida sobre la esfera de Riemann $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Definición 4.1. *Una* función racional $R : \overline{\mathbb{C}} \longrightarrow \overline{\mathbb{C}}$ *es una función de la forma*

$$R(z) = \frac{P(z)}{Q(z)}, \tag{4.4}$$

donde P y Q son polinomios coprimos (sin factores comunes).

El grado de una aplicación racional $R(z) = P(z)/Q(z)$ *está definido como*

$$\text{grado}(R) = \text{máx}\{\text{grado}(P), \text{grado}(Q)\}. \tag{4.5}$$

La n-ésima iteración R^n de una función racional R se define recursivamente a partir de las relaciones $R^2 = R \circ R$ y $R^n = R^{n-1} \circ R$. Nótese que $\text{grado}(R^n) = (\text{grado}(R))^n$ para $\text{grado}(R) \geq 1$.

Las funciones racionales pueden interpretarse como las funciones analíticas de la esfera de Riemann $\overline{\mathbb{C}}$ en sí misma. Informalmente, el grado indica cuántas veces la aplicación R «enrolla» la esfera alrededor de sí misma, y éste es igual al número (contado con multiplicidades) de preimágenes de un punto arbitrario.

El estudio de la dinámica de una función racional R trata de analizar el comportamiento de las órbitas de un punto $z_0 \in \overline{\mathbb{C}}$

$$\text{orb}_R^+(z_0) = \{z_0, z_1 = R(z_0), z_2 = R(z_1) = R^2(z_0), \ldots, z_n = R^n(z_0), \ldots\}.$$

Existen varias preguntas que podemos hacernos sobre una órbita dada $\text{orb}_R^+(z_0)$. Por ejemplo, ¿cuáles son los puntos límites de $\text{orb}_R^+(z_0)$?, o en otras palabras, ¿dónde se acumulan los iterados de un punto z_0 por una función racional?

En la situación ideal, cuando la función racional que estamos estudiando es obtenida al aplicar un método iterativo para aproximar los ceros de un polinomio complejo, los puntos donde se aproximan los iterados de un punto arbitrario por la función de iteración en estudio

deberían corresponder a los ceros del polinomio en cuestión. Lamentablemente como veremos esto no siempre ocurre, pueden existir regiones de área positiva en el plano complejo, de modo que si elegimos una condición inicial en esas regiones, las iteraciones de esos punto no convergen a un cero del polinomio. Por otra parte, si existen dichas regiones nos podemos preguntar ¿qué tipo de conducta puede aparecer en este caso? ¿cuáles son las propiedades topológicas y analíticas del conjunto de puntos límites de $\mathrm{orb}_R^+(z_0)$?

Para poder responder a estos interrogantes sobre el estudio dinámico de una función racional hay que analizar el comportamiento de algunos puntos especialmente relevantes, como son los puntos fijos, los ciclos y los puntos críticos.

4.2.1. Puntos fijos de una aplicación racional

Al igual que se vio para funciones de variable real, los puntos fijos de una función de variable compleja $f(z)$ son las soluciones de la ecuación $f(z) = z$. Además, el infinito también puede ser considerado un punto fijo en $\overline{\mathbb{C}}$. Para el caso particular de una función racional $R(z) = P(z)/Q(z)$, ∞ es un punto fijo si y sólo si $\mathrm{grado}(P) > \mathrm{grado}(Q)$. El resto de puntos fijos son las soluciones de $P(z) - zQ(z) = 0$.

La clasificación de los puntos fijos de una función de variable compleja es similar a la vista para funciones de variable real. En concreto, si z_0 es un punto fijo de una función $R(z)$ con multiplicador asociado $\lambda = R'(z_0)$, entonces z_0 es superatractor si $\lambda = 0$, atractor si $0 < |\lambda| < 1$, indiferente si $|\lambda| = 1$ y repulsor si $|\lambda| > 1$. Además, en el caso de las funciones de variable compleja, los puntos fijos indiferentes se clasifican en:

1. *Racionalmente indiferentes* o *parabólicos* si λ es una raíz de la unidad, esto significa, que existe un número natural n, tal que $\lambda^n = 1$, es decir, λ es de la forma $\lambda = e^{2\pi i\theta}$ con θ racional.

2. *Irracionalmente indiferentes* en otro caso, es decir, $\lambda = e^{2\pi i\theta}$ con θ irracional.

Respecto a la cantidad de puntos fijos que puede poseer una función racional, tenemos el siguiente resultado, cuya demostración es consecuencia directa del teorema fundamental del álgebra.

Teorema 4.1. *Una función racional de grado $d \geq 1$ tiene exactamente $d + 1$ puntos fijos contados con sus multiplicidades.*

En este resultado hemos utilizado el concepto de multiplicidad de un punto fijo de una función racional, que enunciamos a continuación.

Definición 4.2. *Sea $r \in \mathbb{C}$ un punto fijo de una función racional R. Decimos que r tiene* multiplicidad $m \geq 1$ *si, r es una raíz de multiplicidad m de la ecuación $F(z) = R(z) - z = 0$, esto es, $F^{(j)}(r) = 0$ para $j = 0, \ldots, m-1$ y $F^{(m)}(r) \neq 0$. En el caso $m = 1$, decimos que r es un cero simple de R, es decir, $F(r) = 0$ y $F'(r) \neq 0$.*

Como el punto del infinito es un punto de la esfera de Riemann, merece la pena comentar qué se entiende por su multiplicidad. Notemos, en primer lugar que la transformación de Möbius $m(z) = 1/z$, aplica $z = \infty$ en $z = 0$ y viceversa. Definimos

$$S(z) = m \circ R \circ m^{-1}(z), \quad m(z) = \frac{1}{z}. \tag{4.6}$$

Entonces $S(0) = 1/R(\infty)$. Por lo tanto, $z = \infty$ es un punto fijo de R si $z = 0$ lo es de S, es decir, $S(0) = 0$. Además su carácter atractor o repulsor vendrá determinado por el valor

$$S'(0) = \lim_{z\to 0} S'(z).$$

Definición 4.3. *La multiplicidad del punto del infinito $z = \infty$ como punto fijo de una función racional R se define como la multiplicidad del punto $z = 0$ como punto fijo de la función S definida en (4.6).*

Respecto de la carácter del punto $z = \infty$ como punto fijo de una función racional $R = P/Q$, tenemos el siguiente resultado.

Teorema 4.2. *Supongamos que*

$$R(z) = \frac{a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0}{b_m z^m + b_{m-1} z^{m-1} + \cdots + b_0}$$

es una función racional con $n > m$. Entonces $z = \infty$ es un punto fijo de R con el siguiente comportamiento:

1. *Superatractor si $m < n - 1$.*

2. *Atractor si $m \leq n - 1$ y $|a_n| > |b_m|$.*

3. *Repulsor si $m = n - 1$ y $|a_n| > |b_m|$.*

4. *Indiferente si $m = n - 1$ y $|a_n| = |b_m|$.*

Como ya hemos comentado en otras situaciones, si aplicamos el método de Newton a una ecuación no lineal $f(z) = 0$, resulta que los ceros de f son puntos fijos de N_f y recíprocamente. Si la función $f(z)$ es un polinomio de grado $d \geq 2$ con todas sus raíces simples, entonces los $d + 1$ puntos fijos de la función de iteración N_f son precisamente estas d raíces junto con el punto del infinito $z = \infty$. El caso de un polinomio general, se analiza en el siguiente ejemplo.

Ejemplo 4.1. *Analícense los puntos fijos y su multiplicidad de la función de iteración del método de Newton aplicado a un polinomio de grado $d \geq 2$.*

Consideremos un polinomio de la forma

$$f(z)=\prod_{i=1}^{n}(z-\alpha_i)^{m_i},\quad m_1+\cdots+m_n=d,\quad \alpha_i\neq\alpha_j,\ \text{si } i\neq j. \tag{4.7}$$

Entonces

$$\log f(z)=\sum_{i=1}^{n} m_i\log(z-\alpha_i).$$

Derivando la expresión anterior,

$$\frac{f'(z)}{f(z)}=\sum_{i=1}^{n}\frac{m_i}{z-\alpha_i},$$

se llega a que la función de iteración del método de Newton para funciones de la forma (4.7) es

$$N_f(z)=z-\frac{\Pi(z)}{\sum_{i=1}^{n} m_i\Pi_i(z)}=\frac{z\sum_{i=1}^{n} m_i\Pi_i(z)-\Pi(z)}{\sum_{i=1}^{n} m_i\Pi_i(z)}, \tag{4.8}$$

donde hemos definido los siguientes polinomios

$$\Pi(z)=\prod_{i=1}^{n}(z-\alpha_i),\quad \Pi_i(z)=\prod_{\substack{j=1\\ j\neq i}}^{n}(z-\alpha_j).$$

Obsérvese que el denominador de la función racional $N_f(z)$ definida en (4.8) es un polinomio de grado $n-1$, mientras que el numerador es un polinomio de grado n cuyo coeficiente director es $m_1+\cdots+m_n-1=d-1\neq 0$. En consecuencia, la función de iteración del método de Newton tiene $n+1$ puntos fijos que son precisamente las n raíces distintas del polinomio (4.7) junto con el punto del infinito $z=\infty$.

Por último, nótese que el grado de la función de iteración del método de Newton N_f coincide con el grado del polinomio si todas las raíces son simples. Pero si aparece alguna raíz múltiple, entonces la función racional correspondiente N_f tiene grado menor que d. En concreto el grado de N_f es ahora el número n de raíces diferentes del polinomio (4.7).

Ejemplo 4.2. $z=\infty$ *es un punto fijo repulsor para el método de Newton aplicado a un polinomio.*

Si $p(z)$ es un polinomio de grado d, es una comprobación inmediata que $N_p(\infty)=\infty$. Además, como

$$N_p'(\infty)=\frac{d}{d-1}>1,$$

se trata de un punto fijo repulsor.

En los ejemplos anteriores se ha puesto de manifiesto que, para ecuaciones polinómicas, la función racional asociada al método de Newton tiene el número máximo de puntos fijos que establece el teorema 4.1 y que estos puntos fijos son precisamente las raíces distintas del polinomio en cuestión, junto con el punto del infinito. Sin embargo, esta propiedad no es extensible a métodos iterativos para resolver ecuaciones no lineales. En ocasiones, al aplicar algún método iterativo, digamos M, a una función f, la función de iteración obtenida, M_f, puede tener puntos fijos distintos de los ceros de f. A estos puntos fijos los llamaremos *puntos fijos extraños*. Los siguientes ejemplos muestran la aparición de puntos fijos extraños en algunos procesos iterativos conocidos.

Ejemplo 4.3. *La función de iteración asociada al método del punto medio introduce puntos fijos extraños.*

El conocido como método del punto medio para resolver ecuaciones no lineales tiene la siguiente expresión (véase [146]):

$$M_f(z) = z - \frac{f(z)}{f'\left(z - f(z)/(2f'(z))\right)}. \tag{4.9}$$

Cuando aplicamos (4.9) al polinomio $p(z) = z^3 - 1$, obtenemos la función de iteración racional

$$M_p(z) = \frac{z(13z^6 + 22z^3 + 1)}{(5z^3 + 1)^2}$$

que tiene a $z = 0$ como punto fijo que no corresponde con ninguna raíz de $p(z)$.

Además, si r es un cero simple de p, es decir, $p(r) = 0$ y $p'(r) \neq 0$, entonces $M'(r) = 0$ y r es un punto fijo superatractor. De hecho se puede ver que también $M''(r) = 0$ y $M'''(r) \neq 0$ por lo que la convergencia del método en un entorno de estos puntos es cúbica. Sin embargo, para el punto fijo extraño, $z = 0$, tenemos que $M_p'(0) = 1$ y, por tanto, éste es un punto fijo racionalmente indiferente.

Ejemplo 4.4. *La función de iteración asociada al método de Chebyshev (2.8) introduce puntos fijos extraños.*

La expresión del método de Chebyshev para resolver ecuaciones no lineales es la siguiente (véase [146]):

$$Ch_f(z) = z - \left(1 + \frac{L_f(z)}{2}\right)\frac{f(z)}{f'(z)}, \quad L_f(z) = \frac{f(z)f''(z)}{f'(z)^2}.$$

Si aplicamos este método al polinomio $p(z) = 2z^3 + z^2 + z - 1$ obtenemos la función de iteración

$$Ch_p(z) = z^2\frac{30 + 40z + 75z^2 + 54z^3 + 104z^4 + 120z^5}{(1 + 2z + 6z^2)3},$$

que contiene a $z = 0$ como punto fijo aunque no es ninguna raíz de p. Además del $z = 0$, aparecen otros tres puntos fijos extraños, que son las raíces del polinomio $-1 + 23z + 32z^2 + 48z^3$.

Supongamos que se aplica un método numérico concreto para encontrar las raíces de un polinomio $p(z)$ de grado $d \geq 2$, obteniéndose una función de iteración racional que denotamos M_p. En el peor de los casos, si M_p tiene grado $k(d)$, entonces podrían aparecer $k(d)-d$ puntos fijos extraños.

Respecto al problema de aproximar las raíces de una ecuación por un proceso iterativo, la aparición de puntos fijos extraños no sería problemática si éstos son repulsores. Ahora bien, si un punto fijo extraño es atractor, esto daría lugar a la existencia de conjuntos U, de medida positiva, de forma que si $z_0 \in U$, el correspondiente método numérico empezando en z_0 no converge a ninguna solución de la ecuación considerada.

Vamos a estudiar ahora la dinámica asociada a los puntos fijos indiferentes. Comenzamos por el caso

$$R(z) = z - z^{m+1} + O(z^{m+2}), \qquad m \geq 1, \tag{4.10}$$

es decir, $z = 0$ es un punto fijo de multiplicidad $m + 1$. Además, se trata de un punto fijo indiferente pues $M_f'(0) = 1$.

El siguiente resultado describe las cuencas de atracción en este caso.

Teorema 4.3 (de los pétalos)**.** *Sea R una función racional que tiene en $\xi = 0$ un punto fijo racionalmente indiferente como en (4.10). Sean $\omega_1, \ldots, \omega_m$ las raíces m-ésimas de la unidad y sean $\eta_1, \ldots, \eta_m$ las raíces m-ésimas de -1. Entonces existe un radio r_0 y un ángulo θ_0, tal que para $j = 1, \ldots, m$, los sectores S_j y Σ_j, definidos por*

$$S_j = \left\{ z : 0 < \left| \frac{z}{\omega_j} \right| < r_0, \left| \arg\left(\frac{z}{\omega_j} \right) \right| < \theta_0 \right\}$$

y

$$\Sigma_j = \left\{ z : 0 < \left| \frac{z}{\eta_j} \right| < r_0, \left| \arg\left(\frac{z}{\eta_j} \right) \right| < \theta_0 \right\}$$

satisfacen

$$|R(z)| < |z|, \quad \textit{para todo } z \in S_j,$$

y

$$|R(z)| > |z|, \quad \textit{para todo } z \in \Sigma_j.$$

Teorema 4.4. *Sea R una función racional como en el teorema 4.3. Para un $t > 0$ dado y para cada $k = 0, \ldots, m-1$, definimos los pétalos*

$$\Pi_k(t) = \left\{ re^{i\theta} : r < t(1 + \cos(m\theta)), \left| \frac{2k\pi}{m} - \theta \right| < \frac{\pi}{m} \right\}. \tag{4.11}$$

Entonces, para t suficientemente pequeño, R aplica cada pétalo en si mismo. Además, para cada $z \in \Pi_k(t)$, los iterados de z por R convergen a 0.

Se puede dar una prueba análoga a estos resultados para el caso

$$R(z) = z - az^{m+1} + O(z^{m+}), \qquad a \neq 0 \text{ y } m \geq 1, \tag{4.12}$$

con $a \neq 1$. Para más información, véanse por ejemplo [16], [26], [73], [102] o [103].

4.2.2. Ciclos en una función racional

Otro de los fenómenos más frecuentes en el estudio dinámico de una función racional es la aparición de ciclos. Para el caso de funciones de variable compleja, la definición y propiedades básicas de los ciclos son meras generalizaciones de lo que se vio en el capítulo 1 para funciones reales.

En el caso de funciones racionales asociadas a funciones de iteración, aparte de tener puntos fijos extraños, pueden tener ciclos de período mayor que 1. Es obvio que la órbita de un punto inicial que está en un ciclo no convergerá a un cero de la función asociada a la función de iteración.

Ejemplo 4.5 (Smale, [139]). *El método de Newton aplicado al polinomio $p(z) = z^3 - 2z + 2$ tiene un 2-ciclo de la forma $\{0, 1\}$.*

En efecto, la función de iteración asociada a este problema es

$$N_p(z) = \frac{2(z^3 - 1)}{3z^2 - 2}.$$

Una comprobación inmediata nos muestra que $N_p(0) = 1$ y $N_p(1) = 0$.

El carácter atractor, repulsor o indiferente de un n-ciclo viene dado por el valor de su multiplicador asociado. Recordemos que si $\alpha = \{z_0, R(z_0), \dots, R^{(n-1)}(z_0)\}$ es un n-ciclo de R, su multiplicador $\lambda = \lambda(\alpha)$ se define como $\lambda(\alpha) = (R^n)'(z_0)$.

Obsérvese que, por la regla de la cadena, se tiene que

$$(R^n)'(z_0) = \prod_{j=0}^{n-1} R'(R^j(z_0)),$$

y el valor $\lambda(\alpha)$ sólo depende de α, y no del punto particular elegido sobre el ciclo.

Dependiendo del multiplicador asociado, un n-ciclo se clasifica como:

1. Superatractor si $\lambda = 0$.
2. Atractor si $0 < |\lambda| < 1$.
3. Repulsor si $|\lambda| > 1$.
4. Indiferente si $|\lambda| = 1$.

Así, el 2-ciclo que aparece en el ejemplo 4.5 es superatractor ya que

$$(N_p^2)'(0) = N_p'(1)N_p'(0) = 0.$$

Sea M un método iterativo para encontrar raíces de una ecuación no lineal. Al aplicarlo a un polinomio complejo $p(z)$, se cumple que los ceros de $p(z)$ son puntos fijos atractores o superatractores. Sea r uno de tales ceros y supongamos que la primera derivada no nula de M en r es la de orden $k \geq 1$. Entonces, podemos escribir el desarrollo de Taylor de M alrededor de r como sigue

$$M(z) = r + \frac{M^{(k)}(r)}{k!}(z-r)^k + O((z-r)^{k+1}).$$

Entonces, para valores de z cercanos a r, se tiene que

$$M(z) - r \approx \frac{M^{(k)}(r)}{k!}(z-r)^k.$$

En otras palabras, si comenzamos las iteraciones con z_0, y sabemos que $\operatorname{orb}^+(z_0)$ converge a r, llamando $z_{n+1} = M(z_n)$ se tiene que

$$z_{n+1} - r \approx \frac{M^{(k)}(r)}{k!}(z_n - r)^k,$$

es decir, localmente tenemos un buen comportamiento de los iterados por M para puntos de partida «próximos» a la solución del problema, con convergencia de orden k. Pero ¿qué ocurre si elegimos un punto de partida arbitrario? ¿Se puede dar una descripción global de la conducta de los iterados por un método iterativo para encontrar raíces? Para el método de Newton, en la actualidad existe una descripción bastante general para el problema de aproximar las raíces de un polinomio complejo, en particular tenemos los resultados de Hubbard-Schleicher-Sutherland [70] y de Schleicher [134], que expondremos más adelante.

Los siguientes teoremas describen la conducta local de las iteraciones en vecindades de órbitas periódicas atractoras y superatractoras, respectivamente, de una función racional $R(z)$. Para $r > 0$, denotamos por D_r el disco abierto de centro en el origen y radio r, es decir, $D_r = \{z \in \mathbb{C} : |z| < r\}$.

Teorema 4.5 (Königs, [85]). *Si z_0 pertenece a un n-ciclo atractor de una función racional $R(z)$, con multiplicador $\lambda = (R^n)'(z_0)$, que satisface $0 < |\lambda| < 1$, entonces existe una vecindad U de z_0 y un homeomorfismo analítico $\varphi : U \to D_r$ (para algún $r > 0$), único, tal que $\varphi(z_0) = 0$, $\varphi'(z_0) = 1$ y $\varphi(R^n(z)) = \lambda\varphi(z)$ para todo $z \in U$.*

Teorema 4.6 (Böttcher, [23]). *Sea $\operatorname{orb}^+(z_0)$ un n-ciclo superatractor de una función racional $R(z)$. Supongamos que $(R^n)^{(k)}(z_0) \neq 0$, y que*

$$(R^n)'(z_0) = (R^n)''(z_0) = \cdots = (R^n)^{(k-1)}(z_0) = 0.$$

Entonces existe una vecindad U de z_0 y un homeomorfismo analítico $\varphi : U \to D_r$ (para algún $r > 0$) tal que $\varphi(z_0) = 0$, $\varphi'(z_0) = 1$, y $\varphi(R^n(z)) = (\varphi(z))^k$, para todo $z \in U$.

Para la demostración de ambos teoremas el lector puede consultar [16], [21], [22] o [102].

Usando esos dos resultados, podemos definir la *cuenca de atracción* de un punto fijo (super)atractor como sigue. Sea ξ un punto fijo (super)atractor de una función racional R, entonces existe un disco abierto $D_r(\xi)$ de radio $r > 0$ y centro en ξ, tal que para cada $z_0 \in D_r(\xi)$, los iterados $R^n(z_0)$ están definido para todo $n \in \mathbb{N}$, están contenidos en $D_r(\xi)$, y convergen a ξ cuando $n \to \infty$. El conjunto

$$\mathcal{B}(\xi) = \bigcup_{n \geq 0} R^{-n}(D_r(\xi)) \tag{4.13}$$

está formado por todos los puntos en el plano complejo extendido que por iteraciones de R convergen a ξ. En otras palabras, $\mathcal{B}(\xi) = \{z \in \overline{\mathbb{C}} : R^n(z) \to \xi \text{ cuando } n \to \infty\}$. La *cuenca de atracción inmediata*, $\mathcal{B}^*(\xi)$, es la componente conexa de $\mathcal{B}(\xi)$ que contiene a ξ.

La cuenca de atracción de un punto fijo (super)atractor es el conjunto abierto maximal con la propiedad que si elegimos un punto w en $\mathcal{B}(\xi)$, sus iterados $R^n(w)$ convergen a ξ cuando $n \to \infty$.

Si $\alpha = \{z_0, R(z_0), \ldots, R^{(n-1)}(z_0)\}$ es un n-ciclo (super)atractor, entonces su cuenca de atracción es

$$\mathcal{B}(\alpha) = \bigcup_{j=0}^{n-1} R^j(\mathcal{B}(z_0)).$$

Sean $D_j(\alpha)$ las componentes conexas de $\mathcal{B}(\alpha)$ que contienen los puntos

$$z_0, R(z_0), \ldots, R^{(n-1)}(z_0),$$

respectivamente. Llamamos a $D(\alpha) = \bigcup_{j=0}^{n-1} D_j(\alpha)$ la cuenca de atracción inmediata de α, aunque este conjunto no es conexo.

4.2.3. Puntos críticos de una función racional

Dada una función racional $R : \overline{\mathbb{C}} \longrightarrow \overline{\mathbb{C}}$ de grado d, una propiedad importante de la misma es que, excepto a lo sumo en una cantidad finita de puntos $w \in \overline{\mathbb{C}}$, el conjunto $R^{-1}(w) = \{z \in \overline{\mathbb{C}} : R(z) = w\}$ está formado por d elementos distintos. Para verlo, aplicamos el teorema fundamental del álgebra al polinomio $p_w(z) = P(z) - wQ(z)$. Notemos que el grado de p_w es exactamente d, excepto para $w = 0$ y, a lo sumo, para otro valor w para el cual se puede anular el coeficiente principal de p_w. Ahora, si $p_w(z) = 0$ y $p'_w(z) = 0$, es decir,

$$\frac{P(z)}{Q(z)} = w = \frac{P'(z)}{Q'(z)},$$

obtenemos la ecuación $P'(z)Q(z) - P(z)Q'(z) = 0$, que tiene sólo una cantidad finita de soluciones. En otras palabras, excepto posiblemente para una cantidad finita de valores de w, el polinomio $p_w(z)$ no tiene raíces múltiples, por lo tanto tiene d soluciones distintas.

Estas observaciones nos conducen directamente a uno de los conceptos fundamentales en el estudio de la dinámica de funciones racionales, como es el de *punto crítico*.

Definición 4.4. *Sea R una función racional de grado d. Un punto $w \in \overline{\mathbb{C}}$ para el cual la cardinalidad de $R^{-1}(w)$ es menor que d es llamado un* valor crítico *de R. Un punto $z \in R^{-1}(w)$ que es una raíz de $R(z) - w$ con multiplicidad mayor que 1, es llamado un* punto crítico *de R.*

Ejemplo 4.6. *Encuentre los puntos críticos de $R(z) = (z-a)/(z-b)^2$.*

Dado $\omega \in \mathbb{C}$, la ecuación $R(z) = \omega$ tiene dos soluciones distintas, excepto si se anula el discriminante de la ecuación polinómica

$$\omega z^2 - (2b\omega + 1)z + a + \omega b^2.$$

Esto ocurre cuando $\omega = 1/(4(a-b))$ que es, por tanto, un valor crítico. El punto crítico asociado es

$$z = \frac{2b\omega + 1}{2\omega} = 2a - b.$$

Pero, además, si $\omega = \infty$, $R^{-1}(\omega) = b$, por lo que $\omega = \infty$ es otro valor crítico, en este caso, asociado al punto crítico $z = b$.

Intuitivamente, si z es un punto crítico de R, entonces R deja de ser inyectiva en un entorno de z. Es fácil ver que el conjunto de puntos críticos de $R(z) = P(z)/Q(z)$ está formado por los ceros de $R'(z)$, los polos de $R(z)$ con multiplicidad mayor que 1, es decir, los ceros de $Q(z)$ de multiplicidad mayor que 1, y posiblemente, ∞.

Como la dinámica de las transformaciones de Möbius, es decir, funciones racionales de grado uno, es a la vez bien comprendida y elemental (véase [2]), en lo que sigue nos centraremos sólo en la dinámica de funciones racionales R con grado$(R) \geq 2$.

Un cálculo directo usando la regla de la cadena nos permite probar el siguiente resultado.

Teorema 4.7. *Sea $C = C(R)$ el conjunto de puntos críticos de una función racional R. Entonces*

(a) El conjunto de puntos críticos de R^n es

$$C(R^n) = C \cup R^{-1}(C) \cup \cdots \cup R^{-n}(C).$$

(b) El conjunto de valores críticos de R^n es

$$R(C) \cup R^2(C) \cup \cdots \cup R^n(C).$$

El siguiente teorema relaciona las cuencas de atracción y los n-ciclos atractores.

Teorema 4.8 (Fatou y Julia)**.** *La cuenca de atracción inmediata de un ciclo (super)atractor, contiene al menos un punto crítico.*

Este resultado es fundamental, pues nos dice que para determinar los ciclos (super)atractores, debemos estudiar las iteraciones de los puntos críticos de la función de iteración es cuestión.

Teorema 4.9. *Una función racional de grado $d \geq 2$ tiene $2d-2$ puntos críticos contados con su multiplicidad.*

El teorema de Shishikura (1987) da una cota sobre la cantidad máxima de ciclos (super)atractores o indiferentes que puede tener una función racional. Su demostración puede verse en [16], [26] o [102].

Teorema 4.10 (Shishikura, [138])**.** *Una función racional $R : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ de grado d tiene a los más $2d-2$ ciclos, los cuales pueden ser (super)atractores o indiferentes.*

El siguiente resultado muestra que la cota máxima para el número de ciclos puede ser alcanzada.

Teorema 4.11 (Hurley, [71])**.** *Para cada $d \geq 2$ existe un polinomio $p(z)$ de grado d, con coeficientes reales, tal que el método de Newton N_p tiene $2d-2$ ciclos atractores en el plano complejo, es decir, posee el número maximal de ciclos atractores que una función racional de grado d puede tener.*

4.3. Los conjuntos de Fatou y Julia

Existen varias formas de comenzar las exposiciones de la teoría de P. Fatou y G. Julia (1919 y 1918). Adoptamos en estas notas la de Fatou [51], [52]. Éste se basa en el concepto de *familia normal* debido a Montel (véase [2]).

Definición 4.5. *Una familia Γ de funciones meromorfas definidas en un dominio $U \subset \overline{\mathbb{C}}$*

$$\Gamma = \{f_i : U \to \overline{\mathbb{C}} : f_i \text{ meromorfa}\}$$

es normal *si cada sucesión $(f_n)_{n\in\mathbb{N}}$ de elementos de Γ tiene una subsucesión $(f_{n_k})_{k\in\mathbb{N}}$ que converge uniformemente sobre cada subconjunto compacto de U.*

Recordemos que una familia Γ de funciones de un espacio métrico X en un espacio métrico Y es *equicontinua* si, para cada $\varepsilon > 0$ dado, existe $\delta > 0$ tal que

$$d(x_1, x_2) < \delta \text{ implica } d(f(x_1), f(x_2)) < \varepsilon, \text{ para toda } f \in \Gamma.$$

Considerando la métrica cordal o la esférica en $\overline{\mathbb{C}}$, se tiene el siguiente resultado.

Teorema 4.12. *Una familia de funciones meromorfas Γ de U en $\overline{\mathbb{C}}$ es normal si y sólo si es equicontinua sobre todos los subconjuntos compactos de U.*

El siguiente teorema nos proporciona otro criterio para determinar cuándo una familia de funciones es normal.

Teorema 4.13 (Montel). *Sea $\Gamma = \{f : U \to \overline{\mathbb{C}}\}$ una familia de funciones analíticas. Si $\bigcup_{f\in\Gamma} f(U)$ omite tres puntos, entonces Γ es normal, esto es, si existen tres puntos α, β, γ tales que*

$$\bigcup_{f\in\Gamma} f(U) \subset \overline{\mathbb{C}} - \{\alpha, \beta, \gamma\},$$

la familia $\mathcal{F}$ es normal.

Ahora nos centraremos en la familia de iterados $\{R^n : n = 0, 1, 2, 3, \ldots\}$ de una función racional $R : \overline{\mathbb{C}} \longrightarrow \overline{\mathbb{C}}$. En este caso, equicontinuidad significa que iteraciones de puntos próximos no divergen.

Definición 4.6. *Un punto $z \in \overline{\mathbb{C}}$ pertenece al* conjunto de Fatou $\mathcal{F}(R)$ *(también llamado* dominio de normalidad*) si existe una vecindad U de z tal que la familia de iterados*

$$\Gamma = \{R^n : U \to \overline{\mathbb{C}} : n = 0, 1, 2, 3, \ldots\}$$

es normal en U.

El *conjunto de Julia* de R, denotado por $\mathcal{J}(R)$ o simplemente por $\mathcal{J}$, cuando no exista peligro de confusión, es el complemento del conjunto de Fatou, esto es, $\mathcal{J}(R) = \overline{\mathbb{C}} - \mathcal{F}(R)$.

Es claro, a partir de su definición, que $\mathcal{F}(R)$ es abierto y en consecuencia $\mathcal{J}(R)$ es cerrado. Un dominio en el conjunto de Fatou es una componente conexa abierta del conjunto de Fatou.

Del teorema de Montel 4.13, se sigue el siguiente resultado.

Teorema 4.14 (Montel). *Si $z \in \mathcal{J}(R)$ y U es una vecindad de z, entonces $\{R^m(U)\}_{m\in\mathbb{N}}$ cubre todo $\overline{\mathbb{C}}$, excepto a lo más dos puntos.*

Definición 4.7. *Sea U un conjunto abierto no vacío de $\overline{\mathbb{C}}$. Definimos el conjunto de* puntos omitidos *de la aplicación racional $R|_U$ como el conjunto*

$$E_U = \overline{\mathbb{C}} - \bigcup_{n\geq 0} R^n(U).$$

Del teorema 4.14 se tiene el siguiente resultado.

Corolario 4.15. *Sea $z \in \mathcal{J}(R)$ y sea U una vecindad abierta de z. Entonces E_U contiene a lo más dos puntos.*

Teorema 4.16. *El conjunto de puntos omitidos de una función racional R es independiente del punto $z \in \mathcal{J}$ usado para definirlo. Por lo tanto lo podemos denotar simplemente por E_R.*

A continuación, veremos algunas propiedades de los conjuntos de Julia y de Fatou asociados a una función racional $R : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ de grado mayor o igual que 2.

Teorema 4.17. *El conjunto de Julia, $\mathcal{J}(R)$, es no vacío.*

Demostración. Si $\mathcal{J}(R) = \emptyset$ entonces $\mathcal{F}(R) = \overline{\mathbb{C}}$. Luego la familia de iterados $\{R^n : n = 0, 1, 2, \ldots\}$ es normal sobre $\overline{\mathbb{C}}$, y en consecuencia existe una subsucesión $\{R^{n_k}\}$ que converge uniformemente sobre $\overline{\mathbb{C}}$ a una función analítica G. Como G es analítica sobre $\overline{\mathbb{C}}$, es de hecho una función racional, luego $\text{grado}(G) < \infty$. Ahora como $\text{grado}(R^{n_k}) = (\text{grado}(R))^{n_k} \to \infty$ cuando $k \to \infty$. Esto es una contradicción, pues se debe tener que $\text{grado}(G) = \lim_{k\to\infty} \text{grado}(R^{n_k})$. □

Teorema 4.18. *Supongamos que $z_0 \in \overline{\mathbb{C}}$ está sobre un ciclo. Si este ciclo es (super)atractor, entonces está contenido en el conjunto de Fatou, y si es repulsor está contenido en el conjunto de Julia de R.*

Sea M un método iterativo para aproximar soluciones de una ecuación $f(z) = 0$. Una propiedad fundamental que debe tener M es que los ceros de $f(z)$ son punto fijos (super)atractores de M_f, función de iteración obtenida al aplicar M a f. Tenemos así el siguiente resultado.

Teorema 4.19. *Sea M un método iterativo para aproximar ceros de una función $f(z)$. Denotemos por M_f la función de iteración obtenida al aplicar M a f. Entonces los ceros de f están contenidos en el conjunto de Fatou $\mathcal{F}(M_f)$ de M_f.*

Teorema 4.20. *Si z_r es un punto en un ciclo repulsor, entonces*

$$\mathcal{J}(R) = \text{clausura}\{z \in \overline{\mathbb{C}} : R^n(z) = z_r,\ n \in \mathbb{N}\}.$$

Esta proposición nos da un algoritmo para representar gráficamente el conjunto de Julia. Para ello basta encontrar un punto fijo o un punto en un ciclo repulsor y considerar sus preimágenes. Computacionalmente este algoritmo es lento, pero para polinomios de grado pequeño es efectivo.

Teorema 4.21 (Teorema fundamental de Fatou y Julia)**.** *Los ciclos repulsores son densos en $\mathcal{J}(R)$, es decir,*

$$\mathcal{J}(R) = \text{clausura}\{z \in \overline{\mathbb{C}} : z \textit{ pertenece a un ciclo repulsor de } R\}.$$

En particular, existe una cantidad infinita de ciclos repulsores y cada $z \in \mathcal{J}(R)$ es obtenido como límite de puntos en ciclos repulsores.

Esta propiedad es particularmente interesante, pues nos dice que, debido a los errores computacionales (por muy pequeños que ésos sean), al iterar un punto de partida que está sobre el conjunto de Julia de R se obtiene una órbita que tenderá a «alejarse» del conjunto de Julia. En particular, si éste tiene medida de Lebesgue cero, entonces, lo más probable es

que después de un número pequeño de iterados, las siguientes iteraciones estén en el conjunto de Fatou, donde tenemos esperanza de convergencia a alguna de las raíces.

Como una observación interesante, podemos comentar que G. Julia comienza su memoria [72] focalizando su atención sobre la clausura del conjunto de puntos periódicos repulsores, y muestra que su complemento es la unión de dominios sobre el cual la familia de iterados $\{R^n : n \geq 0\}$ es normal.

Teorema 4.22. *El conjunto de Julia de una función racional R es un conjunto perfecto, es decir, $\mathcal{J}(R)$ es cerrado y no tiene puntos aislados.*

El teorema siguiente nos dice que mediante una transformación afín podemos transformar las raíces de un polinomio p sin modificar cualitativamente la dinámica del método de Newton.

Teorema 4.23 (Reescalamiento)**.** *Sea $T(z) = \alpha z + \beta$, con $\alpha \neq 0$, y sea $q(z) = p(T(z)) = p \circ T(z)$. Entonces*

$$T \circ N_q \circ T^{-1} = N_p,$$

esto es, T es una conjugación entre N_p y N_q.

Demostración. Notemos que $p = q \circ T^{-1}$ y que $\alpha^{-1} q'(T^{-1}(z)) = p'(z)$. Ahora,

$$\begin{aligned} T \circ N_q \circ T^{-1}(z) &= T(N_q(T^{-1}(z))) \\ &= T\left(T^{-1}(z) - \frac{q(T^{-1}(z))}{q'(T^{-1}(z))}\right) \\ &= \alpha\left(T^{-1}(z) - \frac{q(T^{-1}(z))}{q'(T^{-1}(z))}\right) + \beta \\ &= \alpha\left(\frac{z}{\alpha} - \frac{\beta}{\alpha}\right) - \frac{q(T^{-1}(z))}{\alpha^{-1} q'(T^{-1}(z))} + \beta \\ &= z - \frac{p(z)}{p'(z)} = N_p(z). \end{aligned}$$

Esto completa la prueba del teorema. □

El teorema 4.23 es válido para muchos métodos, entre los cuales podemos citar los métodos de Halley, Chebyshev o Whittaker, entre otros.

De forma más general, podemos enunciar el siguiente resultado (véase [82], por ejemplo) que establece que la dinámica de dos funciones racionales conjugadas por una transformada de Möbius es equivalente. De hecho, las cuencas de atracción y sus correspondientes fronteras (conjuntos de Julia) asociadas a una función racional son conformemente equivalentes a los de la otra.

Teorema 4.24. *Sea R una aplicación racional y M una transformada de Möbius. Definimos una nueva aplicación racional $S = M \circ R \circ M^{-1}$. Entonces*

$$\mathcal{F}(S) = M(\mathcal{F}(R)) \quad y \quad \mathcal{J}(S) = M(\mathcal{J}(R)).$$

Una aplicación de este tipo de ideas nos conduce al siguiente resultado, que caracteriza el comportamiento dinámico del método de Newton cuando se aplica a polinomios cuadráticos.

Teorema 4.25 (Cayley, Schröder). *Sea $p(z)$ un polinomio cuadrático con sus dos raíces distintas, entonces el método de Newton $N_p(z)$ aplicado a p es conjugado a la función $g(z) = z^2$.*

Demostración. Sean α, β con $\alpha \neq \beta$, las raíces de $p(z)$. Tenemos entonces que $p(z) = (z-\alpha)(z-\beta)$, y sin perdida de generalidad, podemos suponer que $0 \leq |\alpha| \leq |\beta|$. Sea $T(z) = (\beta-\alpha)z+\alpha$, entonces $p \circ T(z) = p(T(z)) = (\beta-\alpha)z((\beta-\alpha)z+\alpha-\beta) = (\beta-\alpha)^2 z(z-1)$. Sean $\lambda = \beta - \alpha$ y $q_\lambda(z) = \lambda^2 z(z-1)$, es decir, $q_\lambda = p \circ T$. Por el teorema 4.23, tenemos que $T \circ N_{q_\lambda} \circ T^{-1} = N_p$. Ahora, tomando $q(z) = z(z-1)$, se ve que $N_q = N_{q_\lambda}$. Luego, N_p es conjugada a N_q. Finalmente, es fácil probar que N_q es conjugada a la aplicación $g(z) = z^2$. Esto termina la prueba. □

Teorema 4.26. *Si $\bar{z} \in \mathcal{J}(R)$ entonces*

$$\mathcal{J}(R) = \text{clausura}\{z \in \overline{\mathbb{C}} : R^n(z) = \bar{z},\ \textit{para algún}\ n \in \mathbb{N}\}.$$

Teorema 4.27. *Sea R una función racional, entonces $\mathcal{J}(R)$ y $\mathcal{F}(R)$ son completamente invariantes, es decir*

$$R(\mathcal{J}(R)) = \mathcal{J}(R) = R^{-1}(\mathcal{J}(R)), \quad R(\mathcal{F}(R)) = \mathcal{F}(R) = R^{-1}(\mathcal{F}(R)).$$

Teorema 4.28. *Los conjuntos de Julia de R y de R^m, con $m \in \mathbb{N}$, son el mismo. En otras palabras, $\mathcal{J}(R) = \mathcal{J}(R^m)$.*

Teorema 4.29. *Sea z_a un punto periódico atractor de una función racional R. Entonces $\mathcal{J}(R) = \partial\mathcal{B}(z_a)$ (∂A denota la frontera del conjunto A).*

Esta proposición nos da una algoritmo bastante eficiente para dibujar el conjunto de Julia. Para ello basta encontrar un punto fijo atractor z_a de R y fijando un error $\varepsilon > 0$ pintamos de un color determinado los puntos z en una región acotada (en general un rectángulo) tales que para algún $n \geq 1$, se tiene $|R^n(z) - z_a| < \varepsilon$. Este algoritmo se conoce como *algoritmo de tiempo de escape.* Por otra parte, podemos definir los conjuntos de nivel, $L_k(z_a)$, $i = 1, 2, 3, \ldots$ como sigue: sea $0 < \varepsilon << 1$ (aquí $<<$ significa «bastante menor que») y sea

$$L_0(z_a) = \{z \in \mathbb{C} : |z - z_a| < \varepsilon\},$$

y, para $k = 0, 1, \ldots,$

$$L_{k+1}(z_a) = \{z \in \mathbb{C} - L_k(z_a) : R^k(z) \in L_k(z_a)\}.$$

Se tiene que

$$\partial L_k(z_a) \to \mathcal{J}(R),$$

donde el límite es tomado respecto de la métrica de Hausdorff en $\mathcal{K}(\overline{\mathbb{C}}) = \{K \subset \overline{\mathbb{C}} : K \text{ es compacto}\}$. Una forma de obtener gráficas vistosas del conjunto de Julia con este algoritmo es colorear cada conjunto de nivel L_k con el color k correspondiente módulo P, donde P es el número de colores de la resolución gráfica del particular monitor, en general $P = 16$. En caso que z_a sea periódico, tomamos $|R^{pn} - z_a| < \varepsilon$, donde p es el período de z_a.

Corolario 4.30. *Si $\mathcal{F}(R)$ es no vacío, entonces el conjunto de Julia de R, $\mathcal{J}(R)$, no tiene puntos interiores.*

Demostración. Sea U un dominio abierto contenido en $\mathcal{J}(R)$. Como $\mathcal{J}(R)$ es invariante,

$$\mathcal{J}(R) \supset \bigcup_{n\geq 0} R^n(U) = \overline{\mathbb{C}} - E_R.$$

Además, como $\mathcal{J}(R)$ es cerrado y E_R contiene a lo más dos puntos, se tiene que $\mathcal{J}(R) = \overline{\mathbb{C}}$. □

El resultado anterior junto con el teorema 4.29 tienen aplicaciones directa en el caso de métodos iterativos para aproximar ceros de una función $f(z)$, pues como estamos asumiendo que las raíces correspondan a puntos fijos (super)atractores, estos pertencen al conjunto de Fatou de la correspondiente función racional M_f, por lo tanto se tiene que el conjunto de Fatou, $\mathcal{F}(M_f)$, es no vacío y $\mathcal{J}(M_f) = \partial\mathcal{B}(\alpha)$, donde α es un cero de f, y $\mathcal{J}(M_f)$ no tiene puntos interiores.

Corolario 4.31. *Si D es un dominio, con $D \cap \mathcal{J}(R) = \mathcal{J}^* \neq \emptyset$, entonces existe $m \in \mathbb{N}$ tal que $R^m(\mathcal{J}^*) = \mathcal{J}(R)$.*

Existe un ejemplo clásico, debido a Lattès, [89], en el que se prueba que el conjunto de Julia asociado a la función racional

$$R(z) = \frac{(z^2+1)^2}{4z(z^2-1)}$$

es todo el plano complejo ampliado, $\mathcal{J}(R) = \overline{\mathbb{C}}$. No obstante, este tipo de situaciones no pueden ocurrir para las funciones de iteración usadas para aproximar ceros de una función.

Definición 4.8. *Sea $\xi \in \mathbb{C}$ un punto fijo (que no es un polo) de una función racional R, con multiplicador $\lambda = R'(\xi) \neq 0$. Decimos que R es linealizable en ξ si, existe una vecindad U de ξ y una función analítica h, tal que $h(\xi) = \xi$, es inyectiva en $U \cup R(U)$, y satisface*

$$h \circ R \circ h^{-1}(z) = \xi + \lambda(z - \xi), \quad \text{para todo } z \in h(U). \tag{4.14}$$

Definición 4.9. *Un punto fijo ξ de una función racional R es aislado si existe una vecindad U de ξ que no contiene otros puntos fijos o sobre un ciclo, aparte de ξ.*

Teorema 4.32 (Siegel)**.** *Sea R una función racional de grado $d \geq 2$. Supongamos que ξ es un punto fijo de R con multiplicador $\lambda \neq 0$. Entonces R es linealizable en ξ si y sólo si ξ es un punto fijo aislado de R.*

Definición 4.10. *Un punto fijo irracionalmente indiferente ξ de una función racional R es llamado un* punto de Siegel*, si R es linealizable en ξ, caso contrario es llamado un* punto de Cremer*.*

Teorema 4.33. *Sea ξ un punto fijo irracionalmente indiferente de una función racional R. Entonces R es linealizable en ξ si y sólo si ξ es aislado.*

Definición 4.11. *Sea D una componente del conjunto de Fatou. Decimos que D es periódica si existe $n \geq 1$ tal que $R^n(D) = D$, y decimos que D es preperiódica si existe $k \geq 1$ tal que $R^k(D)$ es periódico.*

El siguiente teorema caracteriza y clasifica las componentes del conjunto de Fatou asociado a una función racional. Su demostración se puede ver en [16] o [21].

Teorema 4.34 (Dominios no errantes de Sullivan, [142])**.** *Sea R una función racional. Entonces todas las componentes del conjunto de Fatou son preperiódicas. Además, sólo existe una cantidad finita de componentes periódicas.*

Sea U una componente periódica del conjunto de Fatou de R, la cual podemos suponer que es fija, entonces U es de uno de los siguientes cinco tipos:

(i) Superatractoras: contiene un punto fijo superatractor.

(ii) Atractoras: contiene un punto fijo atractor.

(iii) Parabólicas (o dominio de Leau): existe un punto fijo parabólico (o racionalmente indiferente) en su frontera.

(iv) Disco de Siegel: contiene un punto fijo irracionalmente indiferente que es un punto de Siegel. En este caso, U es analíticamente equivalente a un disco y la restricción de R a U es analíticamente conjugada a una rotación de ángulo irracional.

(v) Anillo de Herman: contiene un punto fijo irracionalmente indiferente que es un punto de Cremer. En este caso, U es analíticamente equivalente a un anillo y la restricción de R a U es analíticamente conjugada a una rotación de ángulo irracional.

Teorema 4.35. *Sea U una componente del conjunto de Fatou de una función racional R de grado $d \geq 2$, la cual podemos suponer fija. Sea $C = C(R)$ su conjunto de puntos críticos, y*

$$C^+(R) = \bigcup_{n=1}^{\infty} R^n(C).$$

el conjunto postcrítico de R.

(a) Si U es una componente (super)atractora o parabólica, entonces U debe contener al menos un punto crítico.

(b) Si U es un disco de Siegel o un anillo de Herman, entonces la frontera de U está contenida en la clausura de $C^{+}(R)$.

En particular, si R tiene un anillo de Herman, entonces $\mathcal{J}(R)$ no es conexo.

El teorema anterior toma en cuenta todos los puntos sobre ciclos, excepto aquéllos sobre ciclos irracionalmente indiferentes que pertenecen al conjunto de Julia $\mathcal{J}(R)$, esto es, los puntos de Cremer.

El siguiente resultado muestra la conexión entre los puntos de Cremer y los iterados de puntos críticos.

Teorema 4.36. *Todo punto de un ciclo irracionalmente indiferente que está contenido en el conjunto de Julia $\mathcal{J}(R)$ de una función racional R es un punto límite del conjunto $C^{+}(R)$.*

El siguiente resultado establece la conexión entre los puntos críticos de una función racional y los ciclos de componentes del conjunto de Fatou. En concreto, da una cota sobre el número de componentes de Fatou asociadas a una función racional.

Teorema 4.37 (Shishikura)**.** *Sea R una función racional de grado $d \geq 2$. El número total de componentes de Fatou cíclicas, es decir, componentes cíclicas atractoras, superatractoras, parabólicas, discos de Siegel y anillos de Herman, está acotado por $2d-2$.*

Sobre el número de componentes de Fatou, se tiene el siguiente resultado

Teorema 4.38. *El conjunto de Fatou de una función racional R tiene 0, 1, 2 o una cantidad infinita de componentes.*

Corolario 4.39. *Denotemos por M_p a un método iterativo para aproximar raíces de un polinomio $p(z)$. Supongamos que las raíces de $p(z)$ son puntos fijos atractores o superatractores de M_p y que $p(z)$ tiene al menos 3 raíces distintas. Entonces $\mathcal{F}(M_p)$ tiene una cantidad infinita de componentes.*

El recíproco del corolario anterior no es verdadero. Por ejemplo, si consideramos el método de Chebyshev introducido en (2.8),

$$Ch_f(z) = z - \left(1 + \frac{1}{2}L_f(z)\right)\frac{f(z)}{f'(z)} \tag{4.15}$$

aplicado al polinomio cuadrático $p(z) = z^2 - 1$, vemos que el conjunto de Fatou tiene infinitas componentes, como se aprecia en la figura 4.2.

El comportamiento de los iterados sobre una componente de Fatou es descrito por el siguiente teorema.

Teorema 4.40. *Sea U una componente de Fatou de una función racional R de grado $d \geq 2$, que no contiene valores críticos de R. Entonces $R^{-1}(U)$ consiste de d componentes de Fatou $U_1, \ldots, U_d$ tales que R aplica cada U_i homeomórficamente sobre U. Además, U_i es simplemente conexa.*

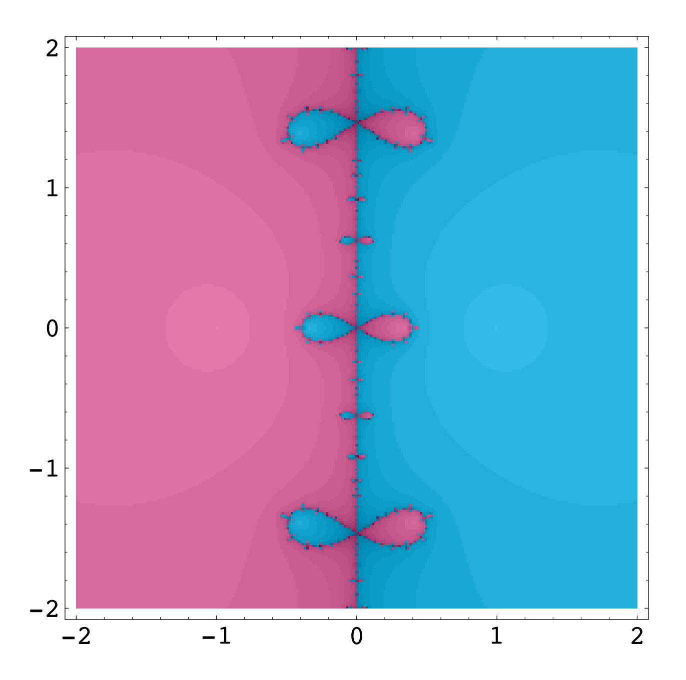


Figura 4.2: Cuencas de atracción del método de Chebyshev aplicado al polinomio $p(z) = z^2 - 1$. En cian aparecen los puntos cuyas órbitas convergen a la raíz $z = 1$ y en magenta los puntos cuyas órbitas convergen a la raíz $z = -1$.

Dada una componente de Fatou arbitraria de una función racional R de grado $d \geq 2$, se tiene que $R^{-1}(U)$ está formado por k componentes distintas $U_1, \ldots, U_k$, on $k \leq d$, las cuales son disjuntas a pares si $k \geq 2$. Además, la restricción de R a U_i es una función racional de grado d_i, con la propiedad que $\sum_{i=1}^{k} d_i = d$. Por lo tanto la preimagen de un punto $z \in \mathbb{C}$ tiene d_i copias en la componente U_i.

Por ejemplo, cuando aplicamos el método de Newton al polinomio $p(z) = z^3 - 1$, obtenemos la función racional

$$R(z) = \frac{2z^3 + 1}{3z^2}.$$

Consideremos la cuenca de atracción inmediata de una de las raíces, digamos $\xi = 1$, la cual denotamos por U_1. Tenemos que $R^{-1}(U_1)$ está formado por dos componentes, U_1 misma y otra que denotamos por U_2, con grado$(R|U_1) = 2$ y grado$(R|U_1) = 1$. Por lo tanto, como la preimagen de un punto en U_1 consiste de 2 preimágenes en U_1 y una en U_2, es claro que ξ tiene multiplicidad 2.

4.4. Algunas propiedades del método de Newton en el plano complejo

Sea $p(z) = a_d z^d + \cdots + a_1 z + a_0$, con $a_d \neq 0$, un polinomio de grado d en $\mathbb{C}$, y sea

$$N_p(z) = z - \frac{p(z)}{p'(z)},$$

su método de Newton.

Veamos algunas propiedades elementales de la dinámica de N_p:

(1) $N_p(z_0) = z_0$ si y sólo si $p(z_0) = 0$, es decir, los puntos fijos de N_p son las raíces de p.

(2) $z = \infty$ es siempre un punto fijo de N_p, y como $N_p'(\infty) = \frac{d}{d-1}$ este punto fijo es repulsor. Por lo tanto, si el método de Newton produce un punto cerca de ∞, sus sucesivas iteraciones se aproximan a una parte compacta de $\mathbb{C}$.

(3) Puesto que $N_p'(z) = \frac{p(z)p''(z)}{(p(z))^2}$, si z_0 es una raíz simple de p, se tiene que $N_p'(z_0) = 0$, esto es, z_0 es un punto fijo superatractor de N_p, lo cual implica que N_p es conjugada a la aplicación $z \to z^k$, para algún $k > 1$ en una vecindad de z_0.

(4) Las raíces múltiples de p son puntos fijos atractores, pero no superatractores, de N_p, pues si z_0 tiene multiplicidad $m > 1$, entonces $N_p'(z_0) = \frac{m-1}{m} < 1$.

(5) Para polinomios genéricos de grado d, esto es, tienen todas sus raíces distintas, el método de Newton es una función racional de grado d. Cuando el polinomio tiene raíces múltiples, N_p tiene grado menor que d.

(6) Los puntos críticos de N_p son las raíces simples y los puntos de inflexión de p. Los puntos críticos de N_p que no son raíces de p los llamaremos *puntos críticos libres*. Recuerde que vimos en la sección anterior que las propiedades del conjunto de Julia de una función analítica en $\overline{\mathbb{C}}$ son frecuentemente determinadas por las órbitas de sus puntos críticos.

(7) Los puntos críticos de p, es decir, las raíces de $p'(z) = 0$, son los polos de N_p.

Las propiedades anteriores caracterizan completamente al método de Newton, es decir, tenemos el siguiente resultado.

Teorema 4.41. *Una función racional $R : \overline{\mathbb{C}} \longrightarrow \overline{\mathbb{C}}$ de grado $d \geq 2$ es la función de Newton de un polinomio de grado mayor o igual que 2 si y sólo si el punto $z = \infty$ es el único punto fijo repulsor y para todos los otros puntos fijos $\xi_1, \ldots, \xi_d \in \mathbb{C}$ existe un número $n_j \in \mathbb{N}$, tal que $R'(\xi_j) = \frac{n_j-1}{n_j} < 1$.*

Este resultado fue probado por G. Saunder en 1984 ([132]). También se le atribuye a J. Head ([63]) en 1987 y a K. Nishizawa y M. Fujimura en 1992 (véase [106]). En particular, ese resultado contiene el caso del método de Newton para polinomios con raíces simples. Para otros métodos, no se tiene una tal resultado, aún en el caso de polinomios con raíces simples.

El primer resultado acerca de la ubicación de los puntos críticos de un polinomio es el teorema clásico de Gauss-Lucas, que enunciamos a continuación.

Teorema 4.42 (Gauss-Lucas, [92])**.** *Los puntos críticos de un polinomio no contante p están contenidos en la envoltura convexa de sus raíces.*

El siguiente teorema es importante para la descripción global de las posible conducta de los iterados por el método de Newton.

Teorema 4.43 (Shishikura, [138]). *Sea R un función racional que posee un único punto fijo repulsor o racionalmente indiferente con multiplicador $\lambda = 1$, entonces $\mathcal{J}(R)$ es conexo.*

Como $z = \infty$ es el único punto fijo repulsor para la función de iteración del método de Newton, N_p, deducimos la siguiente consecuencia.

Corolario 4.44. *Sea $p(z)$ un polinomio complejo, entonces el conjunto de Julia de N_p es conexo.*

Nota 4.1. *Esta propiedad del conjunto de Julia del método de Newton cuando es aplicado a polinomios es una parte fundamental en la demostración del teorema 4.60 de Hubbard, Schleicher y Sutherland y que en esencia dice que el método de Newton es un algoritmo eficiente para el cálculo de raíces de polinomios. Este teorema, junto con el resultado de Schleicher (véase el teorema 4.61), nos permite concluir que el método de Newton es, por tantoo, un algoritmo iterativo.*

4.4.1. El método de Newton para polinomios cuadráticos

Como ya se puso de manifiesto al enunciar el problema de Cayley en los antecedentes de este capítulo, el estudio dinámico del método de Newton aplicado a polinomios de la forma $p(z) = (z - a)(z - b)$ es relativamente sencillo. En el teorema 4.25 se vio que $N_p(z)$ es conjugado con la aplicación $g(z) = z^2$. Veamos ahora una nueva demostración de este resultado, poniendo de manifiesto que el conjunto de Julia $J(N_p)$ es la recta que equidista de los puntos a y b.

Teorema 4.45. *Sea N_p la aplicación de Newton para el polinomio $p(z) = (z-a)(z-b)$, con $a, b \in \mathbb{C}$, $a \neq b$. Entonces N_p es conjugada con la aplicación z^2 mediante la transformada de Möbius $M(z) = (z - a)/(z - b)$. Además $J(N_p)$ es una circunferencia en la esfera compleja que pasa por el punto del infinito, o equivalentemente, $J(N_p)$ es la recta que equidista de los puntos a y b en el plano complejo.*

Demostración. Se puede comprobar por sustitución directa que $R(z) = M \circ N_p \circ M^{-1}(z) = z^2$, aunque el cálculo puede resultar un poco tedioso. Veamos una demostración alternativa que puede resultar más interesante desde el punto de vista matemático. Lo primero, es observar que

$$\begin{array}{ll} N_p(a) = a & M(a) = 0 \\ N_p(b) = b & M(b) = \infty \\ N_p(\infty) = \infty & M(\infty) = 1. \end{array}$$

Entonces, se tiene que:

$$\begin{array}{rcccccccc} R: & z & \to & M^{-1}(z) & \to & N_p(M^{-1}(z)) & \to & M(N_p(M^{-1}(z))) \\ & 0 & \to & a & \to & a & \to & 0 \\ & \infty & \to & b & \to & b & \to & \infty \\ & 1 & \to & \infty & \to & \infty & \to & 1. \end{array}$$

$R(z)$ es una aplicación racional de grado 2 (como $N_p(z)$) que fija el 0, el ∞ y el 1. Además,

$$\begin{aligned} R'(z) &= M'(N_p(M^{-1}(z)))N_p'(M^{-1}(z))(M^{-1})'(z) \\ &= \frac{M'(N_p(M^{-1}(z)))N_p'(M^{-1}(z))}{M'(M^{-1}(z))}. \end{aligned}$$

Como $M'(z) = (a-b)/(z-b)^2$ y $N_p'(z) = L_p(z) = p(z)p''(z)/p'(z)^2$, se tiene que

$$R'(0) = \frac{M'(N_p(a))N_p'(a)}{M'(a)} = N_p'(a) = 0 \quad (M'(a) \neq 0).$$

Por otra parte, como $M'(b) = \infty$,

$$R'(\infty) = \lim_{x \to b} \frac{M'(N_p(x))N_p'(x)}{M'(x)} = \lim_{x \to b} \frac{(x-b)^2}{(N_p(x)-b)^2} N_p'(x) = \lim_{x \to b} \frac{1}{N_p'(x)} = \infty.$$

Así, $R(z)$ tiene una raíz doble en $z = 0$, luego es de la forma

$$R(z) = \frac{z^2}{\alpha z^2 + \beta z + \gamma}.$$

Como $R(\infty) = \infty$, $\alpha = 0$. Como $R'(\infty) = \infty$ y

$$R'(\infty) = \lim_{z \to \infty} \frac{2z(\beta z + \gamma) - \beta z^2}{(\beta z + \gamma)^2} = \frac{1}{\beta},$$

se sigue que $\beta = 0$. Por último, como $R(1) = 1$, $\gamma = 1$ y $R(z) = z^2$. □

4.4.2. El método de Newton para polinomios cúbicos con raíces múltiples

Una ligera variante del estudio realizado en la sección anterior nos permite obtener algunas conclusiones acerca del comportamiento del método de Newton cuando aparecen raíces múltiples. Lo primero observación general que podemos hacer es que cuando se aplica el método de Newton a un polinomio de grado d, la función de iteración resultante tiene grado d cuando las raíces son simples. Sin embargo, cuando las raíces son múltiples, el grado de la función de iteración es menor estrictamente que d. Consideramos en esta sección el caso del polinomio

$$p(z) = (z-a)^2(z-b). \tag{4.16}$$

En este caso, particularmente sencillo, el polinomio tiene una raíz doble y una raíz simple. Analizaremos las dinámicas del método de Newton y estudiaremos cómo son las cuencas de atracción de las raíces de p. Veremos que el comportamiento es totalmente distinto a cuando las raíces del polinomio son simples.

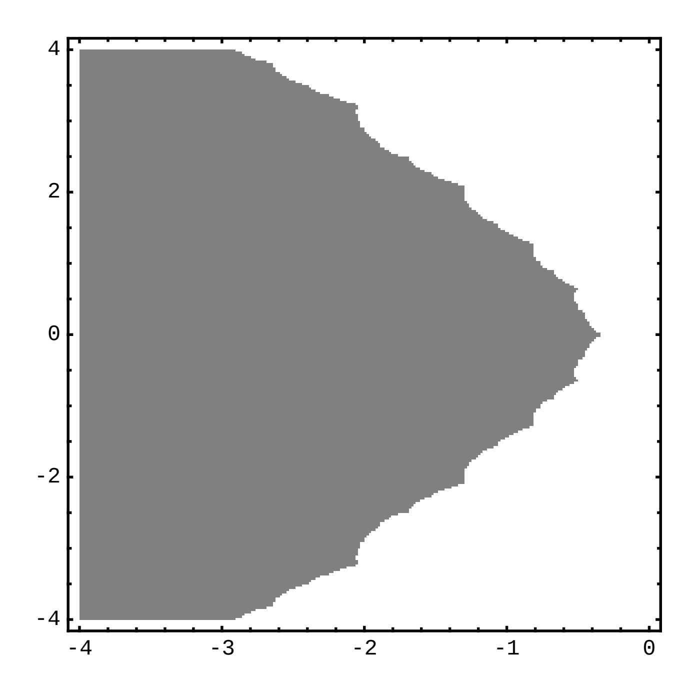


Figura 4.3: Cuencas de atracción de las raíces, $z = -1$ y $z = 1$ para el método de Newton aplicado al polinomio $p(z) = (z-1)^2(z+1)$.

En primer lugar, mediante cambios de variable afines, el estudio del método de Newton aplicado a polinomios de la forma (4.16), puede reducirse al estudio del polinomio $p(z) = (z-1)^2(z+1)$. En este caso, la función de iteración del método de Newton es de la forma

$$N_p(z) = \frac{2z^2+z+1}{1+3z}.$$

Esta función tiene un punto fijo superatractor en $z = -1$ y un punto fijo atractor en $z = 1$, con multiplicador asociado $1/2$. Además, el punto del infinito es un punto fijo repulsor con multiplicador asociado $3/2$. En la figura 4.3 se muestran las cuencas de atracción de las dos raíces, $z = -1$ y $z = 1$. Como se puede apreciar en la figura, la cuenca de atracción de la raíz múltiple, en este caso, $z = 1$, «invade» la cuenca de atracción de la otra raíz, $z = -1$. En este caso, la presencia de dos raíces, una múltiple y otra simple, hace que se pierda la simetría a la que hace referencia el teorema 4.45.

Por otra parte, para el caso de polinomios con raíces múltiples de la forma (4.16) la iteración de Newton, $N_p(z)$, es conjugada mediante la transformada de Möbius $M(z) = (z-a)/(z-b)$ con la aplicación $z(z+1)/2$ definida en el plano complejo ampliado $\hat{\mathbb{C}}$. El correspondiente conjunto de Julia se muestra en la figura 4.4.

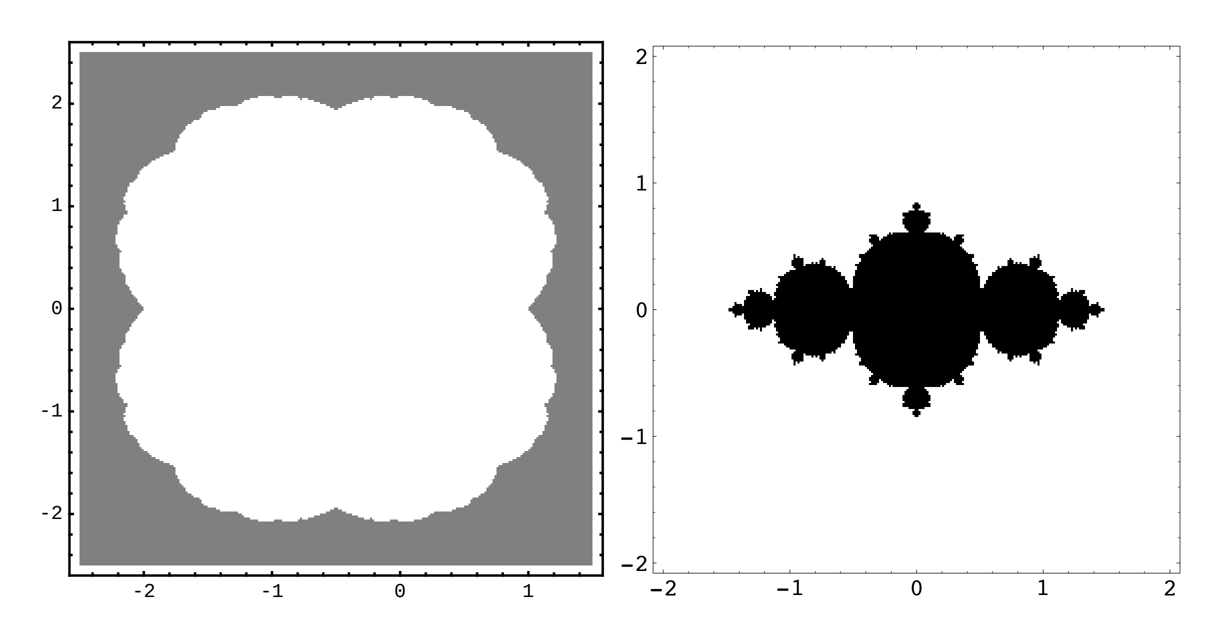


Figura 4.4: Cuencas de atraccción asociadas a las funciones de iteración $z(z+1)/2$ y $z^2-3/4$, relacionadas respectivamente con el método de Newton y el método de Newton para raíces múltiples.

Si se conoce la multiplicidad m de la raíz a aproximar, el conocido como método de Newton para raíces múltiples,

$$N_m(z) = z - m\frac{p(z)}{p'(z)} \tag{4.17}$$

tiene la ventaja de que recupera el orden de convergencia cuadrático al aproximar la raíz múltiple. El estudio de la dinámica del método (4.17) para polinomios de la forma (4.16) lo realizó Gilbert [55]. En ese trabajo, se prueba que la correspondiente función de iteración para polinomios de la forma (4.16) es

$$N_2(z) = z - 2\frac{p(z)}{p'(z)} = \frac{z^2 + az - 2ab}{3z - a - 2b}. \tag{4.18}$$

Esta función racional es conjugada con la aplicación $z^2 - 3/4$ mediante la transformada de Möbius

$$M(z) = \frac{3z + a - 4b}{2(z - a)}.$$

El comportamiento dinámico de la función polinómica $z^2 - 3/4$ es bien conocido (véase [16], por ejemplo). En el segundo gráfico de la figura 4.4 se muestra el conjunto de Julia para $z^2 - 3/4$, que es la frontera de la región de negro. Nótese que en este caso, las raíces a y b del polinomio (4.16) se transforman por M en los puntos

$$M(a) = \infty, \quad M(b) = -\frac{1}{2}.$$

Estos dos puntos ∞ y $-1/2$, junto con el punto $3/2$, son los puntos fijos del polinomio $z^2-3/4$. ∞ es un punto fijo superatractor, $-1/2$ es un punto fijo indiferente y $3/2$ es un punto fijo

repulsor. Deshaciendo los cambios se concluye que el método de Newton para raíces múltiples (4.18) ha transformado la raíz múltiple a en un punto fijo superatractor. Como contrapartida, la otra raíz, b, pasa a ser un punto fijo indiferente. Por último ∞ es un punto fijo repulsor para (4.18).

El segundo gráfico de la figura 4.4 muestra la cuenca de atracción de ∞ como punto fijo de $z^2 - 3/4$. Los puntos de la región de negro convergen al punto fijo indiferente $-1/2$, aunque, en este caso la convergencia es extremadamente lenta.

Otra variante del método de Newton para ecuaciones con raíces múltiples viene dada por

$$\hat{N}_p(z) = z - \frac{1}{1 - L_p(z)} \frac{p(z)}{p'(z)}, \quad L_p(z) = \frac{p(z)p''(z)}{p'(z)^2}. \tag{4.19}$$

$\hat{N}_p(z)$ se obtiene aplicando el método de Newton a la función racional $p(z)/p'(z)$. Para polinomios de la forma (4.16), el método 4.19 es conjugado con la aplicación $-z^2$ mediante la transformada de Möbius $M(z) = (z-a)/(z-b)$. Por lo tanto, su conjunto de Julia es la circunferencia unidad y sus dinámicas son similares al método de Newton para raíces simples (véase el teorema 4.45).

4.4.3. El método de Newton para polinomios cúbicos

Sea $p(z) = a_3z^3 + a_2z^2 + a_1z + a_0$ un polinomio cúbico con sus tres raíces a, b y c distintas, las cuales suponemos ordenadas por sus módulos, es decir, $0 \le |a| \le |b| \le |c|$.

Pongamos $T^{-1}(z) = \alpha z + \beta$, y encontremos los coeficientes α y β de modo que $T^{-1}(a) = 0$ y $T^{-1}(c) = 1$. Tenemos entonces que $\alpha = \frac{1}{c-a}$ y $\beta = -\frac{a}{c-a}$, por lo tanto, $T^{-1}(z) = \frac{z}{c-a} - \frac{a}{c-a}$, y en consecuencia $T(z) = (c-a)z + a$. Aplicando esta transformación T en el teorema de reescalamiento anterior (teorema 4.23), obtenemos

$$q(z) = p \circ T(z) = p((c-a)z + a) = (c-a)^3 z \left(z - \frac{b-a}{c-a} \right) (z-1).$$

Haciendo, $\lambda = (c-a)^3$ y $\rho = (b-a)/(c-a)$, obtenemos

$$q(z) = \lambda^3 z(z-1)(z-\rho).$$

Por otra parte, es fácil ver que si $f(z) = \alpha g(z)$, entonces $N_f(z) = N_g(z)$.

En consecuencia, haciendo

$$p_\rho(z) = z(z-1)(z-\rho) \tag{4.20}$$

y denotando por N_ρ a su correspondiente función de iteración para el método de Newton,

$$N_\rho(z) = z - \frac{z(z-1)(z-\rho)}{3z^2 - 2\rho z - 2z + \rho}, \tag{4.21}$$

tenemos probado el siguiente resultado, que establece que para conocer la dinámica de la método de Newton de un polinomio cúbico debemos conocer la dinámica de la función racional N_ρ, donde $\rho \in \mathbb{C}$ es un parámetro.

Teorema 4.46. *Sea $p(z)$ un polinomio cúbico con sus tres raíces distintas. Entonces, N_p es conjugado topológicamente con N_ρ definida en (4.21).*

Notemos que el caso $\rho = 0$ se reduce al estudio del método de Newton aplicado al polinomio $p_0(z) = z^2(z-1) = z^3-z^2$. Este polinomio tiene en 0 una raíz doble y su comportamiento dinámico es similar al del polinomio que aparece en la figura 4.3.

En este caso, tenemos

$$N_\rho(z) = z - \frac{z^3-(\rho+1)z^2+\rho z}{3z^2-2(\rho+1)z+\rho} = \frac{2z^3-(\rho+1)z^2}{3z^2-2(\rho+1)z+\rho}$$

y

$$N'_\rho(z) = \frac{(z^3-(\rho+1)z^2+\rho z)(6z-2(\rho+1)}{(3z^2-2(\rho+1)z+\rho)^2}.$$

Un estudio sobre de familia fue hecho por Curry, Garnett y Sullivan [37].

En este caso, $N'_\rho(z) = 0$ si y sólo si $p_\rho(z) = 0$ o $p''_\rho(z) = 0$. En consecuencia, el conjunto de puntos críticos de N_ρ está formado por las tres raíces de p_ρ junto con el punto $z = \frac{\rho+1}{3}$. Los puntos críticos de N_ρ que no son raíces de p_ρ se llaman *puntos críticos libres.*

El estudio de las órbitas de los puntos críticos libres da mucha información sobre el comportamiento dinámico de un método. En concreto, para determinar si existen órbitas periódicas atractoras para N_ρ, distintas de las raíces de p_ρ, debemos responder a la pregunta siguiente: ¿para qué valores de ρ, la órbita del punto crítico libre,

$$N_\rho^n\left(\frac{\rho+1}{3}\right)$$

es una órbita periódica atractora?

En la pregunta anterior debemos excluir los casos $\rho = -1$, $\rho = 2$ y $\rho = 1/2$ para los cuales el punto fijo extraño coincide con alguna de las raíces del polinomio p_ρ.

Una manera de responder a la pregunta anterior es colorear el espacio de parámetros $\rho \in \mathbb{C}$ de acuerdo a la convergencia del punto crítico libre $(\rho+1)/3$, tal y como se hace en la figura 4.5. Si la órbita de $(\rho+1)/3$ converge a 0, 1 o ρ, el valor del correspondiente parámetro ρ se colorea en amarillo, cian o magenta respectivamente.

Como se aprecia en la figura 4.5, existen regiones abiertas en el *espacio de parámetros* tales que, si ρ pertenece a estas regiones entonces existen regiones abiertas en el plano complejo de forma que $N_\rho(z)$ definida en (4.21) no converge a ninguna de las raíces del polinomio $p_\rho(z)$ definido en (4.20). Las regiones coloreadas en negro en el espacio de parámetros están formadas por los valores de ρ para los cuales la sucesión

$$N_\rho^n\left(\frac{\rho+1}{3}\right)$$

va a parar a un ciclo atractor.

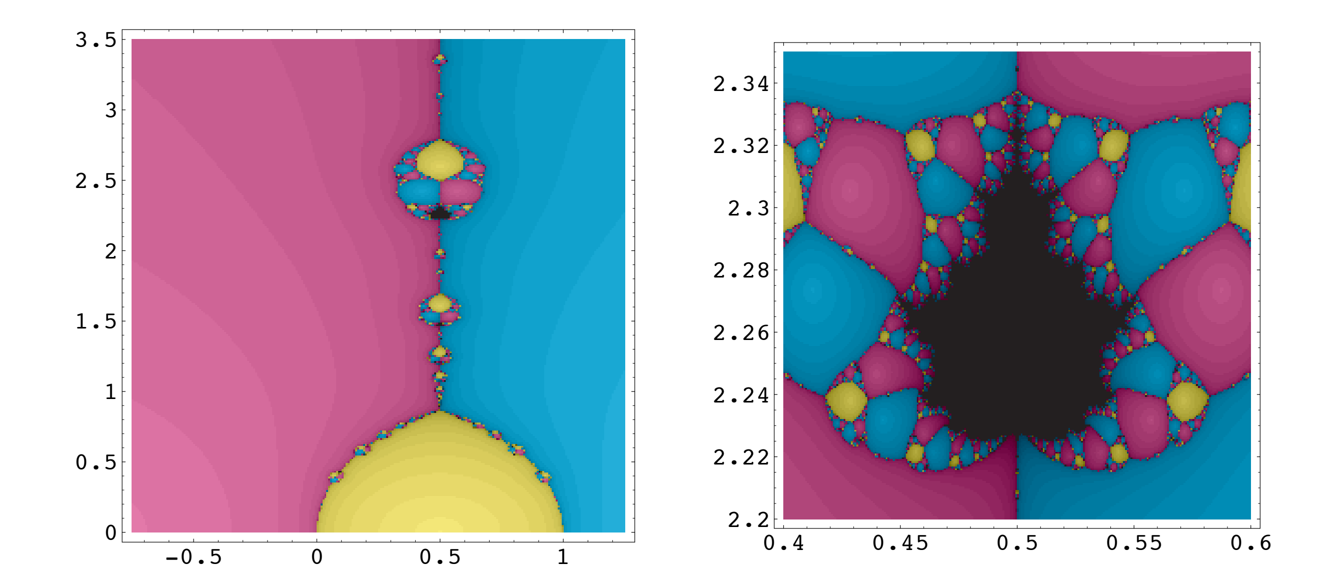


Figura 4.5: Representación gráfica del espacio de parámetros asociado a la función de iteración $N_\rho(z)$ definida en (4.21) y asociada a los polinomios de la forma $p_\rho(z) = z(z-1)(z-\rho)$. La figura de la derecha muestra una ampliación de una zona negra en la que se aprecia un conjunto de tipo Mandelbrot.

La parametrización de los polinomios cúbicos considerada en (4.20) no es la única. Otra parametrización muy habitual (véase [125]) es la siguiente:

$$p_\mu(z) = (z^2-1)(z-\mu), \quad \mu \in \mathbb{C}. \tag{4.22}$$

La correspondiente función de iteración para el método de Newton es

$$N_\mu(z) = \frac{2z^3 - \mu z^2 - \mu}{3z^2 - 2\mu z - 1}. \tag{4.23}$$

En este caso, el punto crítico libre asociado al método de Newton es la única raíz de $p''_\mu(z) = 0$, es decir, $z = \mu/3$. Podemos realizar una reflexiones similares al caso anterior y colorear el espacio de parámetros conforme a la convergencia del punto crítico libre, amarillo, cian o magenta si la órbita de $\mu/3$ converge a μ, 1 o -1 respectivamente, tal y como se muestra en la figura 4.6. De nuevo, las regiones coloreadas en negro en el espacio de parámetros están formadas por los valores de μ para los cuales la sucesión

$$N_\mu^n\left(\frac{\mu}{3}\right)$$

va a parar a un ciclo atractor.

Por último, consideramos otra parametrización muy conocida (véase [118]), como es la siguiente:

$$p_\lambda(z) = z^3 + (\lambda-1)z - \lambda, \quad \lambda \in \mathbb{C}. \tag{4.24}$$

Denotamos $N_\lambda(z)$ a la función de iteración del método de Newton aplicado a los polinomios de la forma (4.24):

$$N_\lambda(z) = \frac{2z^3 + \lambda}{3z^2 + \lambda - 1}. \tag{4.25}$$

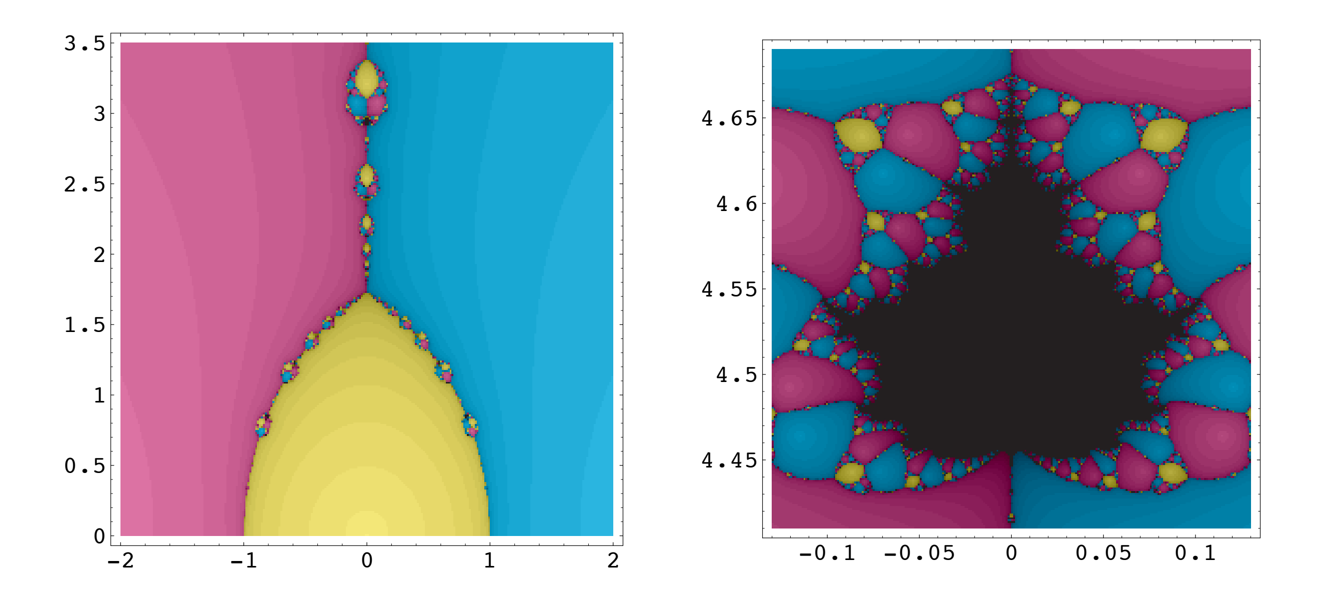


Figura 4.6: Representación gráfica del espacio de parámetros asociado a la función de iteración del método de Newton para los polinomios (4.22). La figura de la derecha muestra una ampliación de una zona negra en la que se aprecia un conjunto de tipo Mandelbrot similar al de la figura 4.5.

En este caso, el punto crítico libre asociado al método de Newton es la única raíz de $p''_\lambda(z) = 0$, es decir, $z = 0$.

Las regiones coloreadas en negro en el espacio de parámetros de la figura 4.7 están formadas por los valores de λ para los cuales la sucesión

$$N_\lambda^n(0)$$

va a parar a un ciclo atractor.

La figura 4.8 muestra las cuencas de atracción del método de Newton para un polinomio $p_\lambda(z)$ definido en (4.24) y tomando λ en una de las zonas negras del espacio de parámetros. Como vemos aparecen «agujeros negros» originados por la presencia de ciclos atractores. En concreto, en este caso se tiene que la órbita del punto crítico libre $z = 0$ es atraída por el 3-ciclo

$$\{1.02169 - 1.04136i, 0.620968 - 0.632698i, -0.00204529 + 0.00527748i\}.$$

4.4.4. El método de Newton para polinomios de grados 4 y 5

Como hemos visto en el apartado anterior, el estudio dinámico del método de Newton aplicado a polinomios de tercer grado se reduce al estudio de una función racional dependiente de un parámetro, en concreto (4.21). Evidentemente, al aumentar el grado de los polinomios también lo hará el número de parámetros involucrados en la correspondiente función racional asociada al método de Newton.

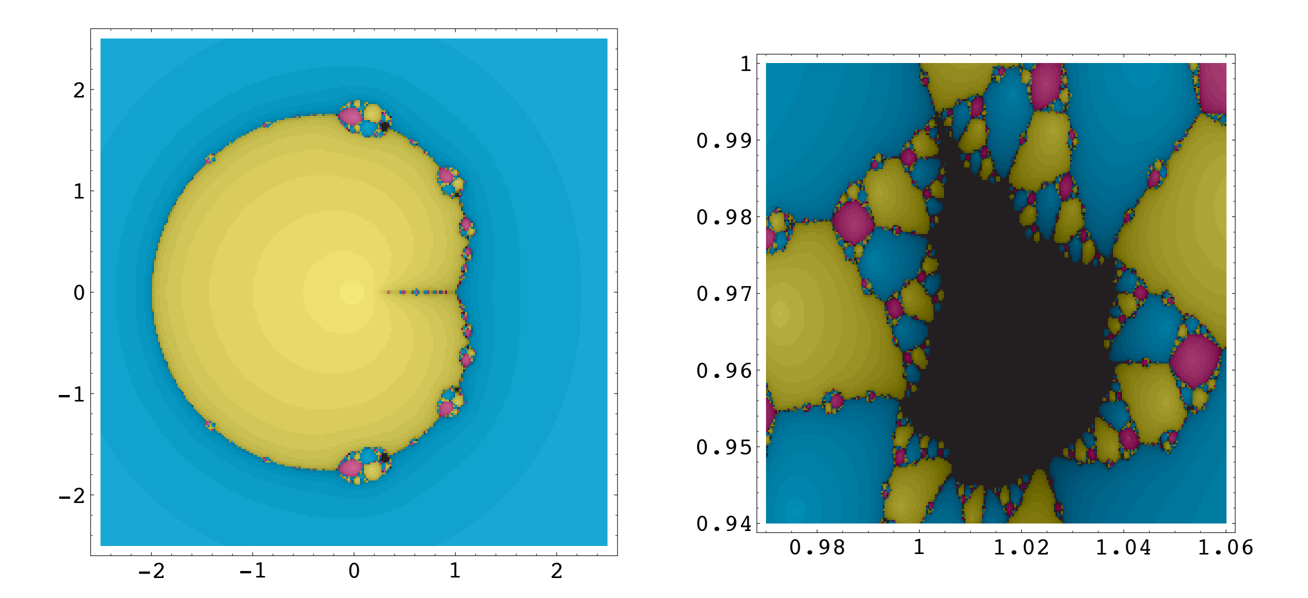


Figura 4.7: Representación gráfica del espacio de parámetros asociado a la función de iteración $N_\lambda(z)$ definida en (4.25) y la correspondiente ampliación mostrando un conjunto de tipo Mandelbrot.

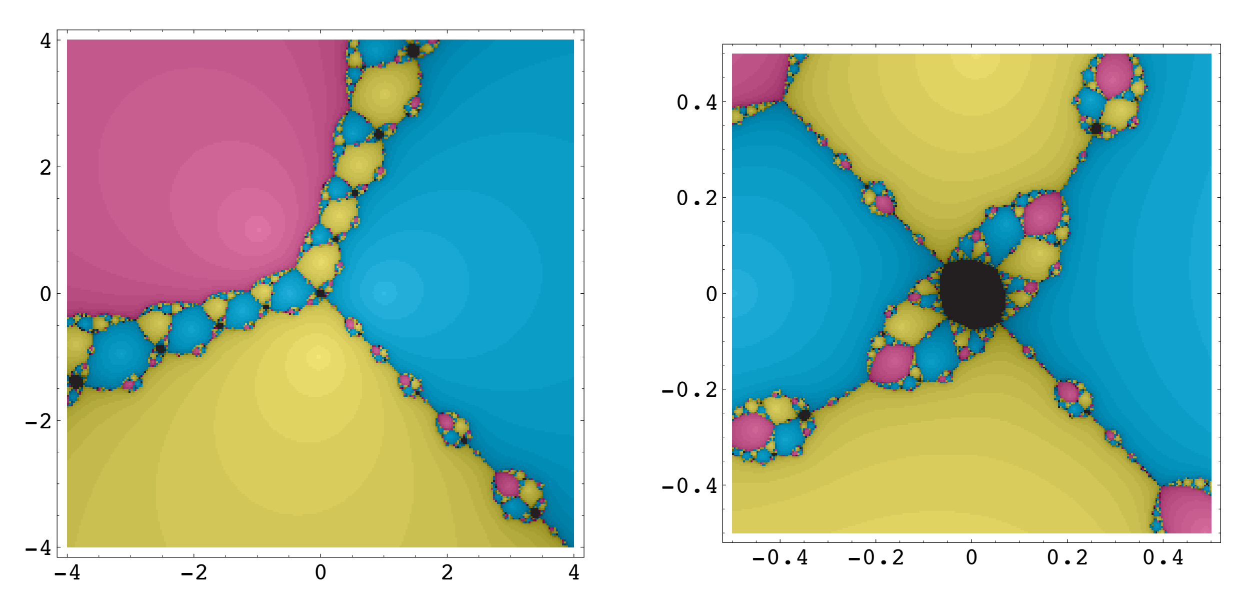


Figura 4.8: Cuencas de atracción del método de Newton aplicado al polinomio $p_\lambda(z) = z^3 + (\lambda - 1)z - \lambda$ con $\lambda = 1.02 + 0.96i$ y una ampliación de la zona negra que se genera entorno al punto crítico libre $z = 0$.

No obstante, existen algunas manipulaciones algebraicas que permiten reducir el número de coeficientes que aparecen en una ecuación polinómica. En concreto, la conocida como *transformación de Tschirnhaus* ([44]) permite transformar la ecuación

$$z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0, \quad n > 2,$$

en otra ecuación polinómica donde no aparecen los términos en z^{n-1} y z^{n-2}, es decir,

$$z^n + b_{n-3}z^{n-3} + \cdots + b_1 z + b_0 = 0, \quad n > 2.$$

El resultado original de Tschirnhaus apareció publicado en *Acta Eruditorum* en 1683. Más adelante, en 1786, E. S. Bring probó que una ecuación polinómica de grado 5 puede reducirse a una del tipo

$$z^5 + az + b = 0.$$

Finalmente, en 1834 G. B. Jerrard demostró que en ecuaciones polinómicas de grado mayor que 3 se puede encontrar una transformación de Tschirnhaus en la que no aparecen los términos en z^{n-1}, z^{n-2} y z^{n-3}, es decir, del tipo

$$z^n + c_{n-4}z^{n-4} + \cdots + c_1 z + c_0 = 0, \quad n > 3.$$

Estas transformaciones se basan en complicadas manipulaciones algebraicas sobre las raíces de la ecuación (véase [155]). Estas manipulaciones no conservan las propiedades dinámicas. En efecto, como se vio en el ejemplo (4.5), el método de Newton aplicado al polinomio $z^3 - 2z + 2$ tiene un 2-ciclo atractor de la forma $\{0, 1\}$. Dicho polinomio puede ser transformado en uno de la forma $p(z) = z^3 - \lambda^3$, $\lambda \in \mathbb{C}$, por la correspondiente transformación de Tschirnhaus. A su vez, el método de Newton aplicado al polinomio anterior,

$$N_p(z) = z - \frac{z^3 - \lambda^3}{3z^2} = \frac{2z^3 + \lambda^3}{3z^2}$$

es conjugado topológicamente, mediante la aplicación afín $h(z) = z/\lambda$, con el método de Newton aplicado al polinomio $q(z) = z^3 - 1$,

$$N_q(z) = z - \frac{z^3 - 1}{3z^2} = \frac{2z^3 + 1}{3z^2}.$$

En efecto,

$$h \circ N_p(z) = \frac{2z^3 + \lambda^3}{3\lambda z^2} = N_q \circ h(z).$$

Pero como se aprecia en las figuras 4.1 y 4.9, el comportamiento dinámico del método de Newton aplicado a los polinomios $q(z) = z^3 - 1$ y $z^3 - 2z + 2$ es muy diferente. De hecho, en el primer caso no aparecen n-ciclos atractores con $n \geq 2$, luego su dinámica no puede ser equivalente a la del método de Newton aplicado al polinomio $z^3 - 2z + 2$, que sí presenta un 2-ciclo atractor.

Para ver que $N_p(z)$ no tiene 2-ciclos atractores, calculamos los puntos fijos de

$$N_p^2(z) = \frac{16z^9 + 51z^6 + 12z^3 + 2}{9z^2\left(2z^3 + 1\right)^2}.$$

Tenemos que $N_p^2(z) = z$ si y sólo si $z^3 = 1$ o $20z^6 + 5z^3 + 2 = 0$. Los 2-ciclos aparecen entre las raíces ξ_j, $j = 1, \ldots, 6$ de la segunda ecuación. Pero en todas ellas se cumple que

$$|N_p'(\xi_j)| \approx 2.45 > 1,$$

luego ningún 2-ciclo es atractor.

Las expresiones simplificadas de polinomios de cuarto y quinto grado que aparecen después de aplicar las transformaciones de Tschirnhaus o de Bring-Jerrard, motivan el estudio del método de Newton para ecuaciones de la forma $z^4 + az + b^4 = 0$ o $z^5 + az + b^5 = 0$ como casos particulares de ecuaciones polinómicas de cuarto y quinto grado respectivamente. En concreto, para la ecuación de quinto grado anterior puede verse un estudio detallado de la dinámica del método de Newton en [8].

4.5. Algoritmos generalmente convergentes

En 1985, Steve Smale [139] planteó la siguiente cuestión sobre la existencia de algoritmos para encontrar los ceros de un polinomio *¿Existen métodos iterativos (punto a punto) dados por una aplicación racional que sean convergentes a las raíces de un polinomio para casi todo punto de partida y para casi todo polinomio?* Podemos pensar en la medida de Lebesgue tanto en el campo complejo como en el espacio $\mathbb{C}^{d+1}$ de los coeficientes de un polinomio de grado d para definir el concepto de *para casi todo* en el sentido de para todos los puntos excepto un conjunto de medida cero.

Dos años más tarde, Curtis McMullen [99] encontró la respuesta a la pregunta anterior. En concreto, McMullen probó lo siguiente:

1. No existen algoritmos generalmente convergentes para polinomios de grado mayor o igual que cuatro.

2. Caracterizó los algoritmos generalmente convergentes para polinomios de grados dos y tres.

Vamos a analizar los resultados dados por McMullen. Para ello, necesitamos las siguientes definiciones:

Definición 4.12. *Sean $\mathbb{P}_d$ el conjunto de polinomios de grado d y $\mathcal{R}at_k$ el conjunto de funciones racionales de grado k definidas de $\hat{\mathbb{C}}$ en $\hat{\mathbb{C}}$. Diremos que $R \in \mathcal{R}at_k$ es convergente para $p \in \mathbb{P}_d$ si la sucesión $\{R^n(z)\}_{n\geq 0}$ converge a una raíz de p para casi todo $z \in \hat{\mathbb{C}}$.*
Un método iterativo dado por R es generalmente convergente *(para polinomios de grado d) si R_p es convergente para casi todo $p \in \mathbb{P}_d$.*

Ya hemos visto que el método de Newton es generalmente convergente para polinomios de grado dos. También es conocido desde los tiempos de Cayley [16] que el método de Newton no es generalmente convergente para polinomios de grado tres o superior, ya que se pueden encontrar polinomios con n-ciclos atractores, $n \geq 2$ y en general polinomios para los cuales el método de Newton no converge a ninguna de las raíces para puntos de partida pertenecientes a un conjunto de medida positiva. A continuación presentamos algunos ejemplos en este sentido.

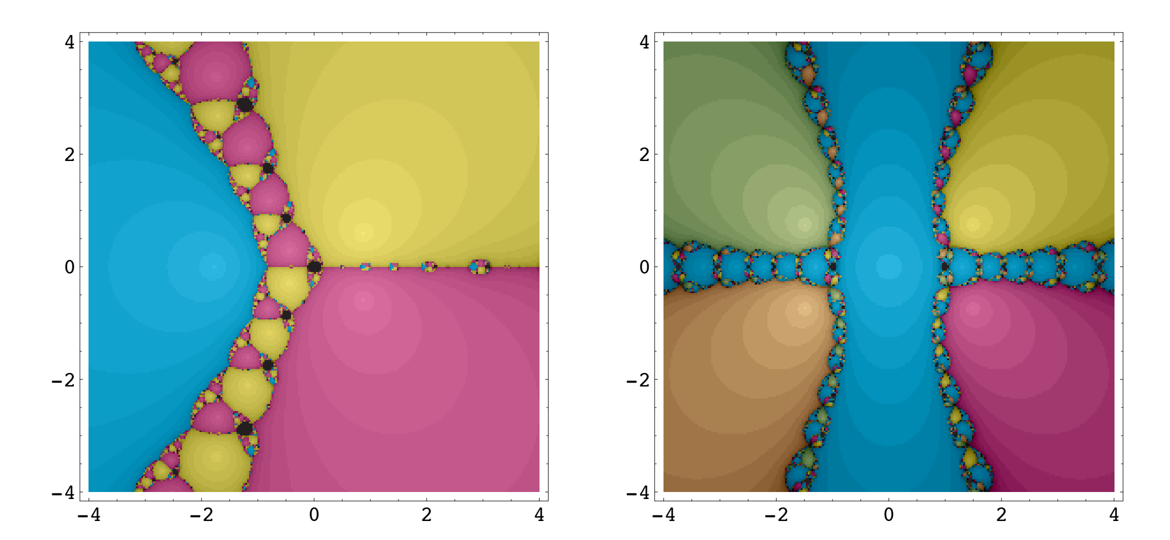


Figura 4.9: Cuencas de atracción del método de Newton aplicado a $p(z) = z^3 - 2z + 2$ y a $p(z) = 3z^5 - 10z^3 + 23z$. En ambos casos se observa la aparición de «agujeros negros» provocados por la presencia de ciclos atractores.

Ejemplo 4.7 (Smale, [139])**.** *Sea $p(z) = z^3 - 2z + 2$. Entonces $\{0, 1\}$ es una órbita periódica superatractora para el método de Newton N_p.*

En efecto, en este caso, tenemos que

$$N_p(z) = \frac{2z^3 - 2}{3z^2 - 2}, \quad \text{y} \quad N_p'(z) = \frac{z(6z^3 - 12z + 12)}{(3z^2 - 2)^2},$$

de donde obtenemos que $N_p'(z) = 0$ si y sólo si $z = 0$ o $p(z) = 0$ (raíces de p), y como $p(0) = 2$ tenemos que $z = 0$ es un punto crítico libre. Notemos ahora que $N_p(0) = 1$ y $N_p(1) = 0$, luego $\{0, 1\}$ es una órbita periódica de N_p y

$$(N_p^2)'(0) = N_p'(N_p(0)) \cdot N_p'(0) = N_p'(1) \cdot N_p'(0) = 0,$$

por lo tanto $\{0, 1\}$ es una órbita periódica superatractora de N_p. En consecuencia, existe un conjunto abierto U de puntos entorno al punto $z = 0$ tal que si $z_0 \in U$, el método de Newton aplicado al polinomio $p(z) = z^3 - 2z + 2$ no converge a ninguna raíz del mismo. De hecho, no es el único conjunto con esta propiedad. Éste y otros conjuntos de puntos de partida para los cuales no hay convergencia a ninguna raíz se muestran como «agujeros negros» en la figura 4.9.

Ejemplo 4.8 (Barna, [11])**.** *Sea $p(z) = 3z^5 - 10z^3 + 23z$. Entonces $\{1, -1\}$ es una órbita periódica superatractora para el método de Newton N_p.*

En este caso, se tiene que los puntos críticos libres son las raíces de $p''(z)$ que no son raíces de $p(z)$, es decir, $z = 1$ y $z = -1$. Como $N_p(1) = -1$, $N_p(-1) = 1$ y

$N'_p(1) = N'_p(-1) = 0$, se tiene que $\{1, -1\}$ es un 2-ciclo superatractor de N_p. En consecuencia, existe un conjunto abierto U de puntos entorno al punto $z = 1$ tal que si $z_0 \in U$, el método de Newton aplicado al polinomio $p(z) = 3z^5 - 10z^3 + 23z$ no converge a ninguna raíz del mismo. De hecho, no es el único conjunto con esta propiedad. Éste y otros conjuntos de puntos de partida para los cuales no hay convergencia a ninguna raíz se muestran como «agujeros negros» en la figura 4.9.

Ejemplo 4.9 (Barna, [11]). *Sea* $p(z) = z^4 - 6z^2 - 11$. *Entonces* $\{1, -1\}$ *es una órbita periódica superatractora para el método de Newton* N_p.

En este caso,

$$N_p(z) = z - \frac{z^4 - 6z^2 - 11}{4z^3 - 12z} = \frac{3z^4 - 6z^2 + 11}{4z^3 - 12z}.$$

De aquí, $N_p(1) = -1$ y $N_p(-1) = 1$, luego $\{1, -1\}$ es una órbita periódica de período 2 para N_p, que además es superatractora ya que $(N_p^2)'(1) = N'_p(-1)N'_p(1) = 0$.

De hecho tenemos más.

Teorema 4.47 (Smale, [139]). *El método de Newton no es generalmente convergente.*

Demostración. Sea $p(z) = \sum_{j=0}^{d} a_j z^j$ un polinomio de grado $d \geq 3$. Fijemos $a_0 = 1$, $a_1 = -1$ y $a_2 = 0$. Tenemos entonces que $p(z) = 1 - z + a_3 z^3 + \cdots + a_d z^d$. Entonces

$$p'(z) = -1 + 3a_3 z^2 + 4a_4 z^3 + \cdots + d a_d z^{d-1},$$

y

$$p''(z) = 6a_3 z + 12a_4 z^2 + \cdots + d(d-1)a_d z^{d-2}.$$

Como $p(0) = 1$, $p'(0) = -1$ y $p''(0) = 0$ se tiene que $N_p(0) = 1$ y $N'_p(0) = 0$. Además, si $N'_p(1) \neq \infty$ entonces $(N_p^2)'(0) = N'_p(1)N'_p(0) = 0$.

Como $N'_p(z) = \frac{p(z)p''(z)}{(p'(z))^2}$, tenemos $N_p(0) = 1$ y $N'_p(0) = 0$, y si $N'_p(1) \neq \infty$ entonces $(N_p^2)'(0) = N'_p(N_p(0))N'_p(0) = N'_p(1)N'_p(0) = 0$. Luego, si $N_p(1) = 0$ y $N'_p(1) \neq \infty$, se tiene que 0 es un punto periódico, de período 2, superatractor para N_p.

Ahora bien, teniendo en cuenta que

$$N_p(z) = z - \frac{1 - z + a_3 z^3 + \cdots + a_d z^d}{-1 + 3a_3 z^2 + 4a_4 z^3 + \cdots + d a_d z^{d-1}},$$

se tiene que $N_p(1) = 0$ si, y sólo si

$$-1 + 2a_3 + 4a_4 + \cdots + (d-1)a_d = 0. \tag{4.26}$$

La condición $N'_p(1) \neq \infty$ se satisface si $p'(1) \neq 0$, es decir, si

$$-1 + 3a_3 + 4a_4 + \cdots + d a_d \neq 0. \tag{4.27}$$

Estas condiciones se satisfacen en un conjunto abierto y denso de un hiperplano de $\{(a_3, a_4, \ldots, a_d) : a_i \in \mathbb{C}\} = \mathbb{C}^{d-2}$. Lo que termina la prueba. □

Por ejemplo, tomando $d = 3$ en el teorema anterior, tenemos el polinomio

$$p(z) = \frac{1}{2}z^3 - z + 1.$$

Notemos que este polinomio es un múltiplo del polinomio que aparece en el ejemplo 4.7. Por lo tanto, la función de iteración obtenida aplicando el método de Newton a ambos polinomios es la misma.

Para $d = 4$ las ecuaciones (4.26) y (4.27) nos dan

$$\begin{cases} -1 + 2a_3 + 3a_4 & = & 0, \\ -1 + 3a_3 + 4a_4 & \neq & 0. \end{cases}$$

De (4.26) obtenemos $a_4 = (1 - 2a_3)/3$ y reemplazando en (4.27) obtenemos $a_3 \neq -1$. Por lo tanto las ecuaciones (4.26) y (4.27) valen para $a_3 \neq -1$. Haciendo $\lambda = a_3$ obtenemos la siguiente familia uniparámetrica de polinomios de grado 4:

$$p_\lambda(z) = 1 - z + \lambda z^3 + \frac{1 - 2\lambda}{3}z^4,$$

o de forma más simple

$$q_\lambda(z) = 3 - 3z + 3\lambda z^3 + (1 - 2\lambda)z^4 \tag{4.28}$$

la cual tiene a $\{0, 1\}$ como órbita periódica superatractora. La figura 4.10 muestra las cuencas de atracción del método de Newton aplicado a uno de estos polinomios, en concreto para $\lambda = 1$. Las regiones en negro corresponden a la cuenca de atracción de la órbita periódica de período 2, $\{0, 1\}$.

De forma más genérica, podemos encontrar una familia de polinomios de grado $n \geq 3$ tales que el método de Newton asociado, $N_p(z)$, tiene a $\alpha = \{0, 1\}$ como un 2-ciclo superatractor, El siguiente resultado da una manera constructiva de calcular dichos polinomios.

Teorema 4.48 (Kalantari, [73]). *Para cada $n \geq 3$, el método de Newton no es generalmente convergente para el polinomio*

$$p(z) = z^n - (n - 1)z + n - 1. \tag{4.29}$$

Demostración. Seguimos las ideas de Smale [139]. Consideramos un polinomio de grado $n \geq 3$ de la forma $p(z) = z^n + az + b$ para el cual imponemos que $\alpha = \{0, 1\}$ sea un 2-ciclo superatractor, es decir, imponemos

$$N_p(0) = 1, \quad N_p(1) = 0, \quad N'_p(0) = 0 \text{ y } N'_p(1) \neq \infty \tag{4.30}$$

Tenemos $N_p^2(0) = 0$, y por la regla de la cadena

$$(N_p^2)'(0) = N'_p(N_p(0)) = N'_p(1)N'_p(0) = 0.$$

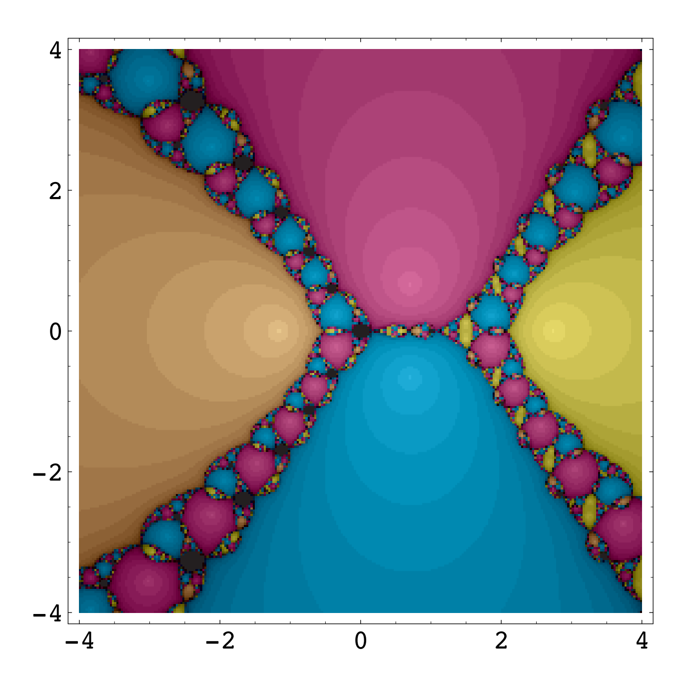


Figura 4.10: Cuencas de atracción del método de Newton aplicado al polinomio $q_1(z) = 3 - 3z + 3z^3 - z^4$ definido en (4.28) cuando $\lambda = 1$.

Ahora, como

$$N_p(z) = z - \frac{p(z)}{p'(z)} = z - \frac{z^n + az + b}{nz^{n-1} + a} = \frac{(n-1)z^n - b}{nz^{n-1} + a},$$

se tiene que $N_p(0) = -b/a$. De la condición $N_p(0) = 1$, se sigue que $b = -a$. Por otra parte, como $N_p(1) = 0$, nos queda $(n-1) - b = 0$, de donde $b = -a = n - 1$. Por lo tanto $p(z) = z^n - (n-1)z + n - 1$. Finalmente, dado que $N_p'(z) = \frac{p(z)p''(z)}{p'(z)^2}$, se sigue que $(N_p^2)'(0) = 0$. □

La figura 4.11 muestra las cuencas de atracción del método de Newton aplicado a uno de los polinomios definidos en (4.29), en concreto para $n = 4$. Las regiones en negro corresponden a la cuenca de atracción de la órbita periódica superatractora $\{0, 1\}$.

La siguiente pregunta que nos planteamos es si podemos construir polinomios para los cuales el método de Newton asociado tenga un ciclo atractor, indiferente parabólico, un disco de Siegel o un punto de Cremer.

Para realizar este estudio, consideremos un polinomio de grado 3, $p(z) = z^3 + a_2z^2 + a_1z + a_0$, e impongamos las condiciones

(1) $N_p(0) = 1$;

(2) $N_p(1) = 0$;

(3) $N_p'(0)N_p'(1) = \lambda$, con $\lambda \in \mathbb{C}$.

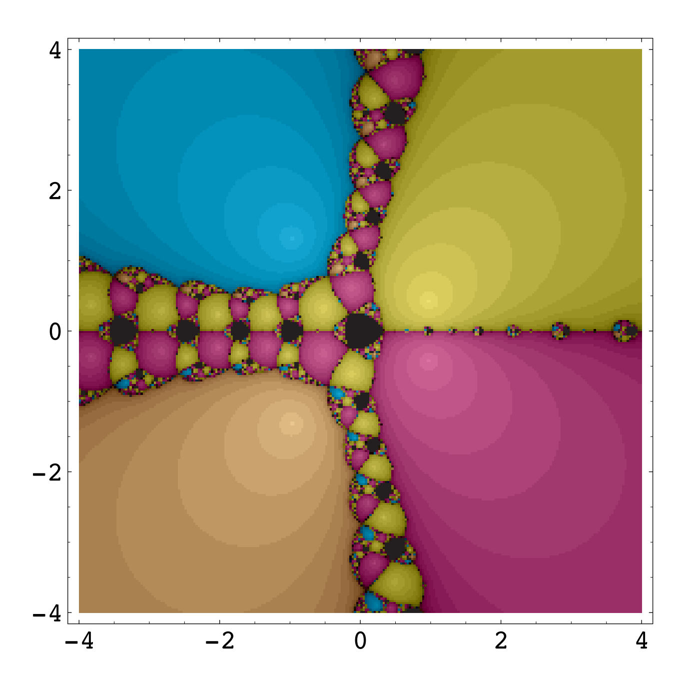


Figura 4.11: Cuencas de atracción del método de Newton aplicado al polinomio $p(z) = z^4 - 3z + 3$ definido en (4.29) para $n = 4$.

Las condiciones $N_p(0) = 1$ y $N_p(1) = 0$ son equivalentes a

$$\frac{p(0)}{p'(0)} = -1 \quad \text{y} \quad \frac{p(1)}{p'(1)} = 1$$

respectivamente. La condición sobre la derivada $N_p'(0) \cdot N_p'(1) = \lambda$ se puede escribir

$$\frac{p(0)p''(0)}{p'(0)^2} \cdot \frac{p(1)p''(1)}{p'(1)^2} = \frac{p(0)}{p'(0)} \frac{p''(0)}{p'(0)} \frac{p(1)}{p'(1)} \frac{p''(1)}{p'(1)} = \lambda,$$

y, por tanto,

$$-\frac{p''(0)}{p'(0)} \frac{p''(1)}{p'(1)} = \lambda.$$

Ahora, para el polinomio $p(z) = z^3 + a_2 z^2 + a_1 z + a_0$, tenemos

$$-1 = \frac{p(0)}{p'(0)} = \frac{a_0}{a_1},$$

es decir, $a_1 = -a_0$. Ahora bien,como

$$1 = \frac{p(1)}{p'(1)} = \frac{1 + a_2 + a_1 + a_0}{3 + 2a_2 + a_1} = \frac{1 + a_2}{3 + 2a_2 - a_0},$$

se sigue que $a_2 = a_0 - 2$. Llamando $\alpha = a_0$ nos quedan las condiciones $a_1 = -\alpha$ y $a_2 = \alpha - 2$. Por lo tanto el polinomio buscado es

$$p_\alpha(z) = z^2 + (\alpha - 2)z^2 - \alpha z + \alpha.$$

Evaluando la condición

$$-\lambda = \frac{p''(0)}{p'(0)} \cdot \frac{p''(1)}{p'(1)}$$

obtenemos que

$$4\frac{(\alpha-2)(\alpha+1)}{-\alpha(\alpha-1} = -\lambda.$$

De aquí se deduce que α debe satisfacer la ecuación cuadrática

$$(4-\lambda)\alpha^2 - (4-\lambda)\alpha - 8 = 0,$$

cuya solución, para $\lambda \neq 4$, es

$$\alpha = \frac{1}{2}\left(1 \pm \sqrt{1 + \frac{32}{4-\lambda}}\right).$$

En definitiva, hemos probado el siguiente resultado.

Teorema 4.49. *Sea $\lambda \in \mathbb{C} - \{4\}$. Entonces el polinomio*

$$p(z) = z^3 + (\alpha-2)z^2 - \alpha z + \alpha$$

con

$$\alpha = \frac{1}{2}\left(1 \pm \sqrt{1 + \frac{32}{4-\lambda}}\right),$$

satisface que $\{0,1\}$ es un 2-ciclo para $N_p(z)$, con multiplicador asociado $(N_p^2)'(0) = \lambda$.

Como aplicación de este resultado, podemos podemos construir polinomios para los cuales el método de Newton asociado tenga un ciclo atractor o un ciclo indiferente parabólico. En concreto, tomando $\lambda = 0$ y $\alpha = 2$ se tiene que el método de Newton aplicado al polinomio $z^3 - 2z + 2$ tiene un 2-ciclo superatractor en $\{0,1\}$ (véase también el ejemplo 4.7). Por otra parte, tomando $\lambda = 1$ y $\alpha = \left(1 + \sqrt{35/3}\right)/2$, la aplicación del teorema anterior nos conduce a un 2-ciclo indiferente para el método de Newton. El comportamiento dinámico en este caso es similar al del ejemplo 4.7), junto a las cuencas de atracción de las tres raíces aparecen las cuencas de atracción del 2-ciclo. La diferencia es que la convergencia hacia el 2-ciclo en el caso indiferente parabólico es mucho más lenta que en el caso atractor. El comportamiento en los casos restantes (discos de Siegel o puntos de Cremer) es bastante más complejo. El lector interesado puede consultar resultados más precisos sobre la dinámica alrededor de estos ciclos parabólicos en [26] o [102].

Una vez vistos varios ejemplos que ponen de manifiesto que el método de Newton no es generalmente convergente para polinomios de grado mayor o igual que 3, pasamos a enunciar los resultados principales de la teoría desarrollada por McMullen.

Teorema 4.50 (McMullen, [99])**.** *Sea T un algoritmo generalmente convergente para polinomios de grado d, en el sentido de la definición 4.12. Entonces T_p y T_q son conjugados por una transformada de Möbius para casi todo p, $q \in \mathbb{P}_d$.*

Demostración. La demostración está basada en el teorema de Thurston sobre puntos críticos de aplicaciones racionales y en el concepto de rigidez de una familia de aplicaciones racionales estable [99, Tma. 2.2]: o bien todos sus miembros son conjugados por una transformada de Möbius, o bien las órbitas de los puntos críticos son finitas. Pensemos en el método de Newton, para el cual,

$$N_p'(z) = \frac{p(z)p''(z)}{p'(z)^2}.$$

La clave está en analizar los puntos críticos que no son raíces de p, es decir, los puntos tales que $p''(z) = 0$ (puntos de inflexión). Esta es la idea que les permite a Roberts y Horgan-Kobelski [125] encontrar «polinomios malos» para el método de Newton. En concreto, para cada $n \geq 2$ encuentran un valor del parámetro λ_n tal que el polinomio $(z^2 - 1)(z - \lambda_n)$ tiene un n-ciclo atractor. □

Corolario 4.51 ([99, Tma. 1.1])**.** *No existen métodos iterativos generalmente convergentes para polinomios de grado mayor o igual que 4.*

Demostración. Supongamos que T es un algoritmo generalmente convergente para $\mathbb{P}_d$. Sean $p, q \in \mathbb{P}_d$, $d \geq 4$, con todas sus raíces simples. Entonces, T_p y T_q son conjugados por una transformada de Möbius M:

$$T_p \circ M = M \circ T_q.$$

Sea s una raíz de q. Entonces s es un punto fijo atractor de T_q, $T_q(s) = s$ y $r = M(s)$ es un punto fijo atractor de T_p. Por la definición de algoritmo generalmente convergente, r es una raíz de p (si $T_p^n(z)$ converge a un punto distinto de una raíz, entonces no puede ser generalmente convergente).

Por lo tanto M transforma las raíces de q en las raíces de p. Dadas dos ternas ordenadas $\{a, b, c\}$ y $\{a', b', c'\}$ existe una y sólo una transformada de Möbius que manda $\{a, b, c\}$ en $\{a', b', c'\}$. Pero no existen transformadas de Möbius que manda $\{a_1, \dots, a_d\}$ en $\{a_1', \dots, a_d'\}$ para $d \geq 4$. Esto prueba que no puede existir un algoritmo generalmente convergente para $d \geq 4$. □

En consecuencia, sólo puede haber algoritmos generalmente convergentes para polinomios de grados dos y tres. McMullen también caracterizó cómo tienen que ser dichos algoritmos. Para ello, necesitaremos la siguiente definición:

Definición 4.13. *Dada una aplicación racional T, definimos el centralizador de T, $C(T)$ como el grupo de las transformadas de Möbius que conmutan con T.*

Corolario 4.52 (Algoritmos generalmente convergentes para polinomios de grado 2, [99, Tma. 1.1])**.** *Un método iterativo generalmente convergente para polinomios de grado 2 se obtiene determinando una aplicación racional T tal que:*

1. *T es convergente para el polinomio $z^2 - 1$.*

2. *El centralizador $C(T)$ contiene a la transformada de Möbius que permuta las raíces cuadradas de la unidad, es decir, $z \to -z$*

El algoritmo es entonces $T_p = A_p T A_p^{-1}$ donde A_p es una transformación afín que manda las raíces cuadradas de la unidad en las raíces del polinomio p.

Corolario 4.53 (Algoritmos generalmente convergentes para polinomios de grado 3, [99, Tma. 1.1])**.** *Un método iterativo generalmente convergente para polinomios de grado 3 se obtiene determinando una aplicación racional T tal que:*

1. *T es convergente para el polinomio $z^3 - 1$.*

2. *El centralizador $C(T)$ contiene a las transformadas de Möbius que permuta las raíces cúbicas de la unidad.*

El algoritmo es entonces $T_p = M_p T M_p^{-1}$ donde M_p es una transformada de Möbius que manda las raíces cúbicas de la unidad en las raíces del polinomio p.

4.5.1. Algoritmos generalmente convergentes para polinomios de segundo grado

Vamos a usar el Corolario 4.52 para determinar la forma de los algoritmos generalmente convergentes para polinomios de segundo grado. Dentro de estos distinguiremos una clase especial que llamaremos superconvergente y estará formada por los algoritmos T_p cuyos puntos críticos son también puntos fijos y, por tanto, coinciden con las raíces de p.

Comenzamos buscando algoritmos superconvergentes dados por una función racional de segundo grado:

$$T(z) = \frac{z^2 + az + b}{cz^2 + dz + e}, \quad (c, d, e) \neq (0, 0, 0).$$

Las condiciones del corolario 4.52 son $T(1) = 1$, $T(-1) = -1$, $T'(1) = 0$, $T'(-1) = 0$ y $T(-z) = -T(z)$. De aquí deducimos que $a = c = e = 0$, $b = 1$ y $d = 2$, con lo que

$$T(z) = \frac{z^2 + 1}{2z} = z - \frac{z^2 - 1}{2z},$$

que no es otra cosa que el método de Newton aplicado al polinomio $z^2 - 1$.

Sean ahora $p(z) = (z - a)(z - b)$, $a \neq b$, $A_p(z) = a + (b - a)(1 - z)/2$ la aplicación afín que manda $1 \to a$ y $-1 \to b$. Entonces,

$$T_p(z) = A_p T A_p^{-1}(z) = \frac{ab - z^2}{a + b - 2z}$$

o equivalentemente,

$$T_p(z) = z - \frac{p(z)}{p'(z)}.$$

Así, el método de Newton es el único algoritmo racional de segundo grado que es superconvergente para polinomios de grado 2.

4.5.2. Algoritmos generalmente convergentes para polinomios de tercer grado

Buscamos la función racional T de menor grado que nos proporcione un algoritmo superconvergente para polinomios de tercer grado. Sean $\omega_k = \exp(2k\pi i/3)$, $k = 0, 1, 2$ las raíces cúbicas de la unidad. De acuerdo con el corolario 4.53, T tiene que cumplir $T(\omega_k) = \omega_k$, $T'(\omega_k) = 0$, $k = 0, 1, 2$. Además T debe conmutar con el conjunto de transformadas de Möbius que permutan las raíces cúbicas de la unidad. Este conjunto es el grupo generado por $M_1(z) = 1/z$ y $M_2(z) = \omega_1 z$. En principio, hay ocho condiciones por lo que parece insuficiente partir de una función racional de tercer grado (siete incógnitas, teniendo en cuenta que podemos normalizar). Así que partimos de

$$T(z) = \frac{z^4 + a_3 z^3 + a_2 z^2 + a_1 z + a_0}{b_4 z^4 + b_3 z^3 + b_2 z^2 + b_1 z + b_0}.$$

Como $T(1/z) = 1/T(z)$, $b_0 = 1$, $b_1 = a_3$, $b_2 = a_2$, $b_3 = a_1$ y $b_4 = a_0$. Además, $T(\omega_1 z) = \omega_1 T(z)$ implica que

$$\frac{\omega_1 z^4 + a_3 z^3 + a_2 \omega_1 z^2 + a_1 \omega_1 z + a_0}{a_0 \omega_1 z^4 + a_1 z^3 + a_2 \omega_1 z^2 + a_3 \omega_1 z + 1} = \frac{\omega_1 z^4 + a_3 \omega_1 z^3 + a_2 \omega_1 z^2 + a_1 \omega_1 z + a_0 \omega_1}{a_0 z^4 + a_1 z^3 + a_2 z^2 + a_3 z + 1},$$

por lo que $a_0 = a_3 = a_2 = 0$ y entonces

$$T(z) = \frac{z^4 + a_1 z}{a_1 z^3 + 1}.$$

Con esto se garantiza además que $T(\omega_k) = \omega_k$, $k = 0, 1, 2$. Falta determinar a_1 para que $T'(\omega_k) = 0$, $k = 0, 1, 2$. Las tres ecuaciones conducen a un único valor $a_1 = 2$, por lo que

$$T(z) = \frac{z(z^3 + 2)}{2z^3 + 1}.$$

Este algoritmo T fue encontrado por McMullen [99]. Se puede comprobar que el orden de convergencia de este algoritmo es tres. Además, T coincide con el método de Halley (2.7) cuando se aplica a $p(z) = z^3 - 1$. En efecto,

$$H_p(z) = z - \left(\frac{2}{2 - L_p(z)}\right) \frac{p(z)}{p'(z)} = z - \frac{3z^3}{2z^3 + 1} \frac{z^3 - 1}{3z^2}$$

$$= z - \frac{z^4 - z}{2z^3 + 1} = \frac{z(z^3 + 2z)}{2z^3 + 1} = T(z).$$

La tentación ahora es extrapolar esta fórmula a un polinomio cualquiera y pensar que el método de Halley y el algoritmo de McMullen coinciden para cualquier polinomio de grado tres. Pero sabemos que el método de Halley no es generalmente convergente para polinomios de tercer grado. Por ejemplo [82], sea $p(z) = 2z^3 + z - 3$, entonces

$$H_p(z) = \frac{12z^5 - 2z^3 + 36z^2 + 3}{24z^4 + 6z^2 + 18z + 1}$$

Cuadro 4.1: Iteraciones del método de Halley aplicado al polinomio $p(z) = 2z^3 + z - 3$ para $z_0 = -1.7$. Se aprecia la presencia de un 2-ciclo atractor.

z_0	-1.7	z_4	-1.707110
z_1	-0.284376	z_5	-0.292898
z_2	-1.707130	z_6	-1.707110
z_3	-0.292918	z_7	-0.292894

y para valores iniciales cercanos a $z_0 = -1.7$ las órbitas son atraídas por un 2-ciclo (véanse la figura 4.12 y el cuadro 4.1).

Luego, en general, el algoritmo de McMullen no puede coincidir con el método de Halley. Veamos, por tanto, cómo es el algoritmo de McMullen.

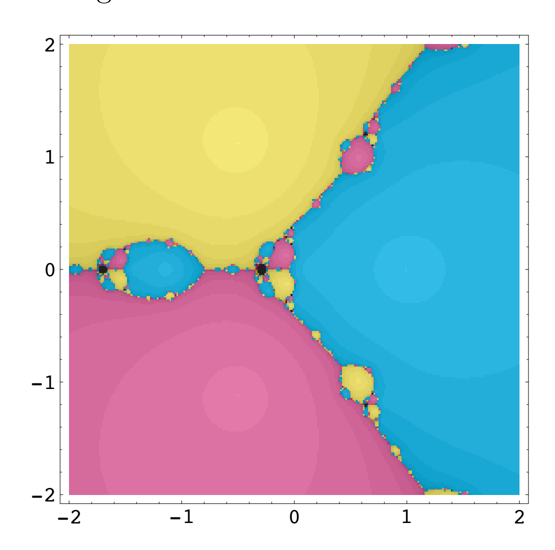


Figura 4.12: Cuencas de atraccción para el método de Halley, $H_p(z)$, aplicado a $p(z) = 2z^3 + z - 3$. Los círculos negros están en la cuenca de atracción de un ciclo atractor no asociado a ninguna de las raíces de p.

Dado un polinomio de tercer grado $z^3 + a_2z^2 + a_1z + a_0$, el cambio de variable afín $z = x - a_2/3$ nos permite transformarlo en uno de la forma

$$x^3 + ax + b,$$

con

$$a = a_1 - \frac{a_2^2}{3}, \quad b = a_0 - \frac{a_1a_2}{3} + \frac{2a_2^3}{27}.$$

Además, si $\zeta \neq 0$ es una raíz de x^3+ax+b, el cambio $x = \zeta\omega$ nos permite escribir la ecuación

$$p_c(\omega) = \omega^3 + (c-1)\omega - c = 0, \quad c \in \mathbb{C}. \tag{4.31}$$

Observemos que como $\zeta^3+a\zeta+b = 0$, $a/\zeta^2+b/\zeta^3 = -1$. Si llamamos $b/\zeta^3 = -c$, entonces $a/\zeta^2 = c - 1$.

Para el desarrollo teórico posterior consideraremos polinomios de la forma (4.31). Las raíces de estos polinomios son:

$$\lambda_0 = 1,\ \lambda_1 = \frac{-1+\sqrt{1-4c}}{2},\ \lambda_2 = \frac{-1-\sqrt{1-4c}}{2}.$$

Con esta particular parametrización de los polinomios de tercer orden, 1 es siempre una raíz y la suma de las tres es cero. Salvo para los casos $c = 1/4$ y $c = -2$, las tres raíces son distintas. Supondremos a partir de ahora que las tres raíces son distintas (esto no afecta al estudio de la convergencia general, pues dejamos de considerar un conjunto de medida cero en el espacio de los polinomios).

Para cada polinomio p_c, el algoritmo de McMullen es entonces

$$T_c(z) = M_c \circ T \circ M_c^{-1}(z)$$

donde M_c es una transformada de Möbius que manda las raíces cúbicas de la unidad a las raíces de p_c, $\lambda_i, i = 0, 1, 2$.

Los puntos fijos de T son ω_0, ω_1, ω_2, 0 e ∞. Observemos además que

$$T'(z) = \frac{2(z^3-1)^2}{(1+2z^3)^2},$$

por lo que las raíces cúbicas de la unidad son puntos críticos de segundo orden (también anulan la segunda derivada) de T.

Como $T_c \circ M_c = M_c \circ T$, los puntos fijos de T_c son $M_c(0)$, $M_c(\infty)$ y $M_c(\omega_k)$, es decir, $M_c(0)$, $M_c(\infty)$ y λ_k.

Sea τ un punto fijo de T tal que $T'(\tau) = 0$. Entonces, $T_c'(M_c(\tau)) = 0$. En efecto,

$$T_c'(M_c(\tau))M_c'(\tau) = M_c'(\tau)T'(\tau).$$

Como $M_c'(\tau) \neq 0$, $T_c'(M_c(\tau)) = 0$.

Si además, $T''(\tau) = 0$, entonces también $T_c''(M_c(\tau)) = 0$. En efecto,

$$T_c''(M_c(\tau))M_c'(\tau)^2 + T_c'(M_c(\tau))M_c''(\tau) = M_c''(\tau)T'(\tau) + M_c'(\tau)T''(\tau) = 0.$$

Como $T_c'(M_c(\tau)) = 0$ y $M_c'(\tau) \neq 0$, $T_c''(M_c(\tau)) = 0$.

En resumen,

$$T_c(z) = z - \frac{p_c(z)(z-M_c(0))(z-M_c(\infty))}{r_c(z)}$$

donde $r_c(z)$ es un polinomio de grado 4 a determinar. Operando se llega a

$$(z-M_c(0))(z-M_c(\infty)) = z^2 - \frac{3c}{c-1}z - \frac{c-1}{3} = \frac{3(c-1)z^2 - 9cz - (c-1)^2}{3(c-1)}.$$

Además, $r_c(z) = (\alpha z + \beta)p_c(z) + (p_c q_c)'(z)$, donde $q_c(z) = 3(c-1)z^2 - 9cz - (c-1)^2$. De aquí se calcula α exigiendo que el algoritmo es de grado 4: $\alpha = -12(c-1)$. Para el cálculo

de β podemos tener en cuenta que $T_c''(1) = T_c''(\lambda_1) = T_c''(\lambda_2) = 0$. Las tres ecuaciones nos conducen a la misma solución $\beta = 18c$.

En definitiva el algoritmo de McMullen para un polinomio de la forma 4.31 es

$$T_c(z) = z - \frac{(z^3 + (c-1)z - c)(3(c-1)z^2 - 9cz - (c-1)^2)}{(3c-1)z^4 - 18cz^3 - 6(c-1)^2z^2 + 6c(c-1)z + 1 - 3c - 6c^2 - c^3}.$$

Para un polinomio de la forma $p(z) = z^3 + az + b$ McMullen [99, Prop.1.2] probó que el algoritmo tiene la forma

$$T_p(z) = z - \frac{(z^3 + az + b)(3az^2 + 9bz - a^2)}{3az^4 + 18bz^3 - 6a^2z^2 - 6abz - 9b^2 - a^3}. \tag{4.32}$$

Como vemos, el algoritmo se expresa en función de los coeficientes de p y, en principio, no admite una expresión en términos de p y sus derivadas. Además, se puede observar que no coincide con el método de Halley aplicado a $p(z) = z^3 + az + b$. En la figura 4.13 se muestran los dominios de atracción para los algoritmos de Halley y McMullen aplicados al polinomio $z^3 + (2i-1)z - 2i$.

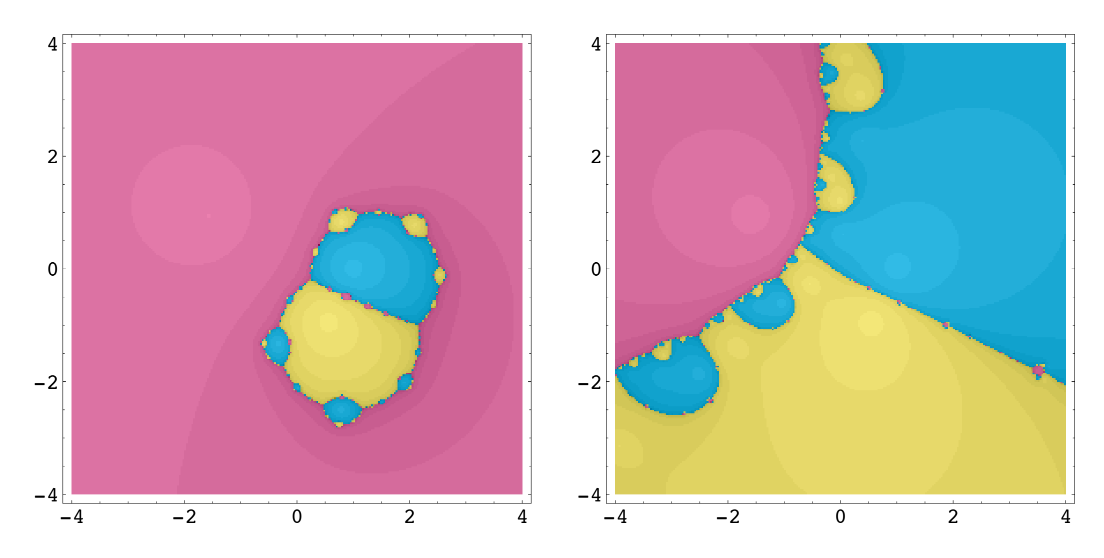


Figura 4.13: Algoritmos de McMullen y de Halley para el polinomio $z^3 + (2i-1)z - 2i$.

El propio McMullen propone una forma alternativa de construir el algoritmo (4.32) como resultado de aplicar el método de Newton a la función racional

$$q(z) = \frac{z^3 + az + b}{3az^2 + 9bz - a^2}.$$

4.5.3. Otros algoritmos generalmente convergentes para polinomios de tercer grado

Aplicando el Corolario 4.53, Jane Hawkins [61] caracterizó las funciones racionales que generan un algoritmo globalmente convergente para polinomios cúbicos.

Teorema 4.54 (Hawkins, [61, Tma. 3.1])**.** *Si R genera un algoritmo globalmente convergente para polinomios cúbicos, entonces existen constantes $a_0, a_1, \ldots, a_k$, con a_0 y a_k distintas de cero, tales que*

$$R(z) = \frac{z(a_0 + a_1 z^3 + \cdots + a_k z^{3k})}{a_k + a_{k-1} z^3 + \cdots + a_0 z^{3k}}.$$

Demostración. Supongamos que $R(z) = p(z)/q(z)$, siendo p y q dos polinomios coprimos (sin factores comunes). La demostración se sigue utilizando únicamente que $R(\omega z) = \omega R(z)$ para ω una raíz cúbica de la unidad (distinta de 1), que $R(1/z) = 1/R(z)$ y que $z = 1$ es un punto fijo de R. El resto de condiciones del Corolario 4.53 servirán, en su caso, para determinar los coeficientes $a_0, \ldots, a_k$. □

Este resultado nos indica que no podemos construir algoritmos generalmente convergente a partir de funciones racionales de cualquier grado. Por ejemplo, para polinomios cúbicos, no existirán algoritmos generalmente convergentes de grados 5 ó 6. Algunos ejemplos de algoritmos globalmente convergentes para polinomios cúbicos son los siguientes:

$$R_4(z) = \frac{z(2 + z^3)}{1 + 2z^3},$$

$$R_7(z) = \frac{z(14 + 35z^3 + 5z^6)}{5 + 35z^3 + 14z^6},$$

$$R_{10}(z) = \frac{z(7 + 42z^3 + 30z^6 + 2z^9)}{2 + 30z^3 + 42z^6 + 7z^9}.$$

El algoritmo R_4 es el que hemos calculado para obtener el algoritmo de McMullen. Se puede comprobar que el orden de convergencia de este algoritmo es tres. R_7 y R_{10} fueron encontrados por Hawkins. R_7 es el único algoritmo de grado 7 cuyo orden de convergencia es mayor que 3. De hecho, tiene orden de convergencia 5. R_{10} es el único algoritmo de grado 10 y orden de convergencia 7.

4.5.4. Conjunto de Julia universal para el algoritmo de McMullen

Para cualquier polinomio cúbico p_c de la forma (4.31), con sus raíces distintas, ya hemos visto que el algoritmo de McMullen T_c es conjugado con

$$T(z) = \frac{z(2 + z^3)}{1 + 2z^3}.$$

Consideramos ahora la transformada de Möbius

$$M(z) = \frac{\sqrt{3}i - (2 + e^{2\pi i/3})z}{e^{2\pi i/3}z - 1}$$

que transforma las raíces cúbicas de la unidad en -1, 1 e ∞ y definimos

$$T_*(z) = M \circ T \circ M^{-1}(z) = \frac{(3 + 6z^2 - z^4)}{8z}.$$

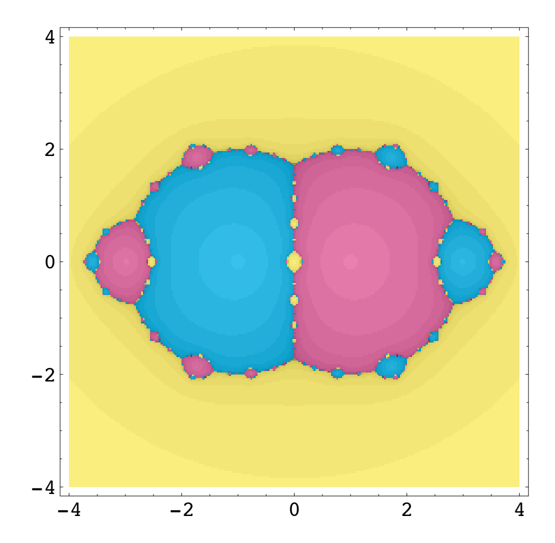


Figura 4.14: Conjunto de Julia universal para un polinomio de tercer grado, obtenido a partir de la aplicación racional $(3 + 6z^2 - z^4)/(8z)$.

El conjunto de Julia para T_* se muestra en la figura 4.14 y lo llamamos conjunto de Julia universal [82] para el método de McMullen. Por tanto, para casi todo polinomio cúbico con las tres raíces distintas, el correspondiente conjunto de Julia para el algoritmo de McMullen es como el de la figura anterior. Ambos son conformemente equivalentes mediante una transformación de Möbius.

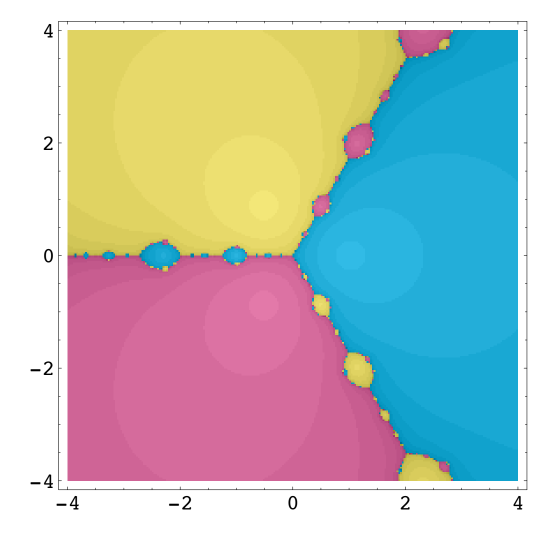


Figura 4.15: Conjunto de Julia para el algoritmo de McMullen aplicado a $z^3 - 1$.

La única excepción es el polinomio $z^3 - 1$. En este caso, el conjunto de Julia tiene una forma diferente (figura 4.15). Esto es debido a que en este caso, para $c = 1$, $T_1(z) = T(z)$, luego ∞ es un punto fijo. Sin embargo para $c \neq 1$, el punto fijo del infinito va a parar a

$$M_c(\infty) = \frac{9c \pm \sqrt{81c^2 + 12(c-1)^3}}{6(c-1)} \neq \infty.$$

En consecuencia, el ∞ deja de ser un punto fijo. Lo normal entonces es que las cuencas de atracción estén acotadas para dos de las tres raíces de p_c, resultando una figura como la de la figura 4.14, salvo un factor de escala. Esto tiene consecuencias interesantes desde el punto de vista numérico, ya que lo que suele ocurrir es que ∞ pertenezca a la cuenca de atracción de alguna de las raíces de p_c y para puntos de partida grandes en módulo se produzca una convergencia rápida hacia esa raíz. Sin embargo, como es sabido, para el método de Newton la convergencia es lineal para valores grandes de $|z|$ y para el método de Halley y para puntos de partida grandes en módulo se podría dar la convergencia a cualquiera de las tres raíces. En el cuadro 4.2 comparamos la convergencia de los métodos de Newton, Halley y McMullen para el polinomio $p_{2i}(z) = z^3 + (2i - 1)z - 2i$ para distintos puntos de partida. Las raíces de p_{2i} son aproximadamente 1, $-1.5643 + 0.93956i$ y $0.5643 - 0.93956i$.

Para terminar este capítulo estudiaremos algunas limitaciones que tienen los algoritmos iterativos para aproximar raíces de polinomios en el plano complejo. Comenzamos con la siguiente definición.

Definición 4.14. *Dado un polinomio $p(z)$ de grado $d \geq 2$. Una función racional $T_p(z)$, de grado $k(d)$, definida en términos de z, $p(z)$ y sus derivadas se llama una* función de iteración *para tal polinomio si, cada raíz de $p(z)$ es un punto fijo (super)atractor de $T_p(z)$. Si esta propiedad vale para todo polinomio $p(z)$ de grado d, decimos que la función racional T definida por $T(p(z)) = T_p(z)$ es un* algoritmo puramente iterativo.

Por ejemplo, el método de Newton

$$N : p(z) \to N_p(z) = z - \frac{p(z)}{p'(z)}$$

es un algoritmo puramente iterativo.

Una pregunta inmediata es ¿cuán bueno es un método iterativo para aproximar las raíces de un polinomio?, o más específicamente, ¿el método iterativo en cuestión converge para casi toda condición inicial? Esto nos lleva a la siguiente teorema, que fue probado por Barna para el caso especial de polinomios con todas sus raíces reales.

Teorema 4.55 (Barna, [12])**.** *Sea $p(z)$ un polinomio con todos sus raíces reales, entonces los puntos de inflexión de $p(z)$ están contenidos en las cuencas de atracción inmediatas para N_p de las raíces $\xi_1 \ldots, \xi_d$ de $p(z)$. Además, excepto por un conjunto de Cantor C de números reales, cada número real converge por iteraciones por N_p a una raíz de $p(z)$.*

Para el método de Newton aplicado a un polinomio cúbico, este puede fallar debido a la existencia de un ciclo atractor. Como ya indicamos el método de Newton aplicado a polinomios no puede tener anillos de Herman, pues su conjunto de Julia es conexo, y consecuentemente toda componente de Fatou es simplemente conexa. La aparición de discos

Cuadro 4.2: Métodos de Newton, Halley y McMullen para $p_{2i}(z) = z^3 + (2i-1)z - 2i$ y $z_0 = 1000 + 100i$, $z_0 = -1000 + 100i$ y $z_0 = -100 - 1000i$.

n	Newton	Halley	McMullen
0	$1000 + 100i$	$1000 + 100i$	$1000 + 100i$
1	$666.667 + 66.6662i$	$500 + 49.9991i$	$-2.40531 + 1.20676i$
2	$444.445 + 44.4434i$	$250.001 + 24.9978i$	$-1.57506 + 0.937464i$
3	$296.297 + 29.6279i$	$125.002 + 12.4955i$	$-1.56432 + 0.939565i$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
19	$0.995525 - 0.0010387i$	1	$-1.56432 + 0.939565i$
20	$1.00002 - 0.0000074i$	1	$-1.56432 + 0.939565i$

n	Newton	Halley	McMullen
0	$-1000 + 100i$	$-1000 + 100i$	$-1000 + 100i$
1	$-666.667 + 66.6662i$	$-500. + 50.0008i$	$-2.3935 + 1.19451i$
2	$-444.445 + 44.4454i$	$-250.001 + 25.002i$	$-1.57461 + 0.937299i$
3	$-296.297 + 29.6312i$	$-125.003 + 12.5041i$	$-1.56432 + 0.939565i$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
19	$-1.5643 + 0.939569i$	$-1.56432 + 0.939565i$	$-1.56432 + 0.939565i$
20	$-1.56432 + 0.939565i$	$-1.56432 + 0.939565i$	$-1.56432 + 0.939565i$

n	Newton	Halley	McMullen
0	$-100 - 1000i$	$-100 - 1000i$	$-100 - 1000i$
1	$-66.6662 - 666.666i$	$-49.9992 - 500i$	$-2.40551 + 1.19343i$
2	$-44.4435 - 444.444i$	$-24.998 - 249.999i$	$-1.57489 + 0.937122i$
3	$-29.6281 - 296.295$	$-12.4959 - 124.997i$	$-1.56432 + 0.939565i$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
19	$0.563728 - 0.937347i$	$0.564322 - 0.939565i$	$-1.56432 + 0.939565i$
20	$0.564326 - 0.93957i$	$0.564322 - 0.939565i$	$-1.56432 + 0.939565i$

de Siegel para métodos iterativos, en especial para el método de Newton, es consecuencia de un resultado de McMullen ([99], [100]).

Una pregunta que surge de inmediato es ¿cómo puede fallar una función de iteración para aproximar raíces para dejar de ser generalmente convergente? Sabemos que la existencia de ciclos atractores nos lleva a una explicación de la pregunta anterior, notemos que la existencia de un ciclo atractor para un elemento $p \in \mathbb{P}_d$, conjunto de polinomios de grado d, implica la existencia de una vecindad abierta $\mathcal{N}_p$ de p en $\mathbb{P}_d$, de modo que cada $q \in \mathcal{N}_p$ falla a converger por la existencia de un ciclo atractor del mismo largo que el de p. Pero ¿es ésta la única

forma de que una función de iteración deje de ser generalmente convergente? Imaginemos que para una función de iteración cuando es aplicada a un polinomio, el conjunto de Julia obtenido tiene medida positiva (ya sabemos que no puede tener puntos interiores), entonces tal función de iteración fallaría a converger sobre un conjunto de medida positiva. Esto no implica que lo mismo ocurra para polinomios próximos. Para ver la existencia de discos de Siegel, argumentamos como sigue. Si $\{T_\lambda\}_\lambda$ es una familia de funciones racionales con un ciclo que cambia de repulsor a atractor cuando λ varía, entonces este ciclo debe contener el centro de Siegel para algún valor del parámetro λ.

Para una función de iteración $T_p(z)$ asociada a un polinomio $p(z)$, tenemos el siguiente resultado.

Teorema 4.56. *Una función de iteración $T_p(z)$ asociada a un polinomio $p(z)$ es generalmente convergente si y sólo si el conjunto de Julia $\mathcal{J}(T_p)$ tiene medida de Lebesgue cero y para todo $z \in \mathcal{F}(T_p)$ se tiene que $\operatorname{orb}^+(z)$ converge a una raíz de $p(z)$. En particular, esto significa que no existen ciclos de componentes de Fatou de largo mayor o igual que dos, ya sean (super)atractores, parabólicos, discos de Siegel o anillos de Herman.*

Definición 4.15. *Decimos que un punto crítico c de una función racional R es preperiódico si existe un menor entero positivo algún $n \geq 1$, tal que $R^n(c)$ está sobre un p-ciclo. Cuando $p = 1$, decimos que c es prefijo.*

Lo anterior significa que n es el primer entero positivo, tal que $c \in R^{-n}(z_0)$, donde z_0 es un punto sobre un p–ciclo.

Teorema 4.57 (McMullen, [100]). *Sea R una función racional. Supongamos que R tiene al menos un ciclo atractor. Si todos los puntos críticos de R están en cuencas de atracción de atractores o son preperiódicos, entonces $\mathcal{J}(R)$ tiene medida de Lebesgue cero. Además, para todo $z \notin \mathcal{J}(R)$ se tiene que sus iterados por R convergen a un ciclo atractor.*

De esto deducimos el siguiente resultado:

Teorema 4.58. *Una función de iteración $T_p(z)$ asociada a un polinomio $p(z)$ es generalmente convergente si sus puntos críticos son preperiódicos o convergen a una raíz de $p(z)$.*

Corolario 4.59. *Sea $N_p(z)$ el método de Newton asociado a un polinomio $p(z)$. Si las raíces de $p''(z)$ son preperiódicas o convergen a una raíz de $p(z)$, entonces N_p es generalmente convergente.*

El siguiente resultado (teorema 1 en [70]) nos muestra que aún cuando el método de Newton no es generalmente convergente, podemos controlar los puntos iniciales de modo a que siempre tengamos convergencia a una raíz. Comenzamos introduciendo la siguiente notación. Sea $\mathcal{P}_d$ el espacio de polinomios de grado d, normalizados de modo que todas sus raíces estén en el disco abierto unitario $\mathbb{D}$ en el plano complejo.

Teorema 4.60 (Hubbard, Schleicher y Sutherland, [70]). *Para cada $d \geq 2$, existe un conjunto $\mathcal{S}_d$ consistiendo de a lo más $1.11d \log^2 d$ puntos en $\mathbb{C}$, con la propiedad que para cada polinomio $p \in \mathcal{P}_d$ y cada una de sus raíces, existe un punto $s \in \mathcal{S}_d$ en la cuenca de atracción de la raíz elegida. Para polinomios cuyas raíces son todas reales, existe un conjunto análogo $\mathcal{S}$ con a lo más $1.3d$ puntos.*

Note que el teorema anterior no hace que el método de Newton sea de hecho un algoritmo, pues fijada una raíz ξ de un polinomio $p \in \mathcal{P}_d$ y un error (tolerancia) $\varepsilon > 0$, no se tienen cotas sobre el número de iterados que debemos calcular para obtener $|\xi - z_n| < \varepsilon$, donde $z_0 \in \mathcal{S}_d$ lo escogemos en la cuenca de atracción de ξ. Este problema fue resuelto por Schleicher en [134], donde caracterizó un conjunto de condiciones iniciales eficientes.

Teorema 4.61 (Schleicher, [134]). *Dado $\varepsilon > 0$, para cada grado d existe un conjunto finito de condiciones iniciales $\mathcal{S}_d \subset \mathbb{C}$ conteniendo $cd(\log d)^2$ puntos con la siguiente propiedad. Dado $p \in \mathcal{P}_d$, entonces para cada raíz ξ de $p(z)$, existe al menos un punto $z \in \mathcal{S}_d$ tal que las iteraciones $N_p^n(z)$ convergen a ξ. Además, el número de iteraciones requerido, M_ε de modo que*

$$|N_p^{M_\varepsilon}(z) - \xi| < \varepsilon$$

depende polinomialmente de d con exponente bajo. El conjunto $\mathcal{S}_d$ puede ser especificado explícitamente.

4.6. Método de Newton para funciones enteras

Sea $f : \mathbb{C} \to \mathbb{C}$ una función entera, es decir, una función holomorfa sobre todo el plano complejo. El método de Newton asociado a f, N_f, es una función meromorfa, es decir, una función holomorfa en todo $\mathbb{C}$ excepto en un conjunto de puntos aislados, llamados polos de la función. La función $N_f : \mathbb{C} \to \overline{\mathbb{C}}$ puede ser extendida a ∞ como una función racional si y sólo si f tiene la forma especial $f(z) = p(z)e^{q(z)}$, donde p y q son polinomios (véase [129, proposición 2.11]). Como en el caso de funciones racionales, se tiene que si $\xi \in \mathbb{C}$ es una raíz de multiplicidad $m \geq 1$ de f, entonces $N_f(\xi) = \xi$ y $N_f'(\xi) = 1 - 1/m$, y recípromente, cada punto fijo finito de N_f es atractor y una raíz de f, y esto de hecho caracteriza las aplicaciones de Newton de funciones enteras (véase [129, proposición 2.8]).

Teorema 4.62 (Rückert y Schleicher, [129]). *Sea $N : \mathbb{C} \to \overline{\mathbb{C}}$ una función meromorfa. Esta es la aplicación de Newton de una función entera $f : \mathbb{C} \to \mathbb{C}$ si y sólo si para cada punto fijo $N(\xi) = \xi \in \mathbb{C}$, existe un número natural $m \in \mathbb{N}$ tal que $N'(\xi) = 1 - 1/m$. En este caso, existe una constante $c \in \mathbb{C} - \{0\}$ tal que*

$$f = c \exp\left(\int \frac{d\zeta}{\zeta - N(\zeta)} \right).$$

Dos funciones enteras f y g tienen la misma aplicación de Newton si y sólo si $f = cg$ para alguna constante $c \in \mathbb{C} - \{0\}$.

Observemos, por el teorema anterior, que una función meromorfa sin puntos fijos es automaticamente la aplicación de Newton de una función entera sin ceros.

Como corolario, se tiene la caracterización de las funciones racionales que son el método de Newton de algún polinomio complejo.

Teorema 4.63. *Una función racional $R : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ de grado $d \geq 2$ es la aplicación de Newton de un polinomio de grado al menos 2 si y sólo si $R(\infty) = \infty$ y para todos los otros puntos fijos $a_1, \ldots, a_d \in \mathbb{C}$, existen $m_1, \ldots, m_d \in \mathbb{N}$ tal que $R'(a_j) = (m_j - 1)/m_j < 1$. Entonces, R es el método de Newton aplicado al polinomio*

$$p(z) = a \prod_{=1}^{d} (z - a_j)^{m_j}$$

para cualquier número complejo $a \neq 0$.

Note que no hemos considerado el caso de una aplicación racional de grado 1 en el teorema anterior, esto es debido al siguiente resultado.

Teorema 4.64. *Sea $f : \mathbb{C} \to \mathbb{C}$ una función entera tal que su función de Newton tiene sólo un punto fijo atractor $\xi \in \mathbb{C}$ con cuenca de atracción inmediata $U = \mathbb{C}$. Entonces existe $d > 0$ y $a \in \mathbb{C}$ tal que $f(z) = a(z - \xi)^d$.*

El siguiente resultado clasifica las funciones racionales que son el método de Newton de funciones enteras.

Teorema 4.65 (Rückert y Schleicher, [129])**.** *Sea $f : \mathbb{C} \to \mathbb{C}$ una función entera. Su aplicación de Newton N_f es una función racional si y sólo si existen polinomios p y q tal que f tiene la forma $f = pe^q$. En este caso, ∞ es un punto fijo repulsor o parabólico.*

Más precisamente, sean $m, n \geq 0$ los grados de p y q, respectivamente. Si $n = 0$ y $m \geq 2$, entonces ∞ es repulsor con multiplicador $\frac{m-1}{m}$. Si $n = 0$ y $m = 1$, entonces N_f es constante. Si $n > 0$, entonces ∞ es parabólico con multiplicador $+1$ y multiplicidad $n + 1 \geq 2$.

Notemos que si N_f tiene una singularidad esencial en ∞, entonces la sucesión de iterados no está definida para puntos $z \in \mathbb{C}$ para los cuales $N_f^n(z) = \infty$ para algún $n \in \mathbb{N}$, y en este caso, $z \in \mathcal{J}(N_f)$ por definición. De hecho, $\mathcal{J}(N_f)$ es la clausura de tales puntos.

Como en el caso de las funciones racionales, cuando f es meromorfa, siempre se tiene que $\mathcal{J}(N_f)$ es no vacío (véase [20, teorema 3]). Por otra parte si f tiene al menos una raíz, entonces $\mathcal{J}(N_f)$ es pequeño en el sentido que es nunca denso y tiene interior vacío (véase [20, lema 3]). En este caso, aparte de las clásicas componentes de Fatou que existen en el caso de funciones racionales, aparece una componente no errante más, llamada *dominio de*

Baker, y pueden aparecer otras componentes errantes. Recordemos que U es una componente errante si $N_f^n(U) \cap N_f^m(U) = \emptyset$, para todo $m \neq n$. Tenemos el siguiente resultado para la clasificación de las componentes del conjunto de Fatou $\mathcal{F}(N_f)$, que podemos encontrar, por ejemplo en [20].

Teorema 4.66. *Sea f una función entera y N_f la función de iteración del método de Newton. Sea U una componente de $\mathcal{F}(N_f)$. Entonces, U es o bien errante o existe $k \in \mathbb{N}$ y una componente periódica V de $\mathcal{F}(N_f)$ tal que $N_f^k(U) \subset V$. En este caso, sea $p \geq 1$ minimal tal que $N_f^p(V) \subset V$, entonces V es exactamente de uno de los siguientes tipos:*

1. *V es una cuenca de atracción, es decir, contiene un único punto periódico atractor z_0 de período p, tal que $N_f^{np}(z) \to z_0$ para todo $z \in V$ cuando $n \to \infty$.*
2. *V es una componente parabólica, esto es, ∂V contiene un único punto periódico parabólico z_0 de período p tal que $N_f^{np}(z) \to z_0$ para todo $z \in V$ cuando $n \to \infty$.*
3. *V es un disco de Siegel, es decir, $N_f^p|_V$ es conformemente conjugado a una rotación irracional del disco unitario.*
4. *V es un anillo de Herman, esto es, $N_f^p|_V$ es conformemente conjugada a una rotación irracional de un anillo de módulo finito.*
5. *V es un dominio de Baker, esto es, $N_f^p(z) \to \infty$ para todo $z \in V$, e ∞ es una singularidad esencial de N_f.*

Notemos que en los casos (2) y (3) del teorema anterior necesariamente se tiene que $p > 1$.

Además, para funciones enteras, se tiene el siguiente resultado que caracteriza las cuencas de atracción de un punto fijo atractor. Su demostración puede verse también en [20].

Teorema 4.67. *Sea $f : \mathbb{C} \to \mathbb{C}$ una función entera no constante. Si ξ es un punto fijo atractor del método de Newton N_f, entonces su cuenca de atracción inmediata es simplemente conexa y no acotada.*

Para más resultados sobre el método de Newton para funciones no racionales, en especial funciones enteras y meromorfas en el plano complejo, recomendamos al lector revisar los artículos [20], [62], [129] y las referencias que allí aparecen. Para finalizar esta sección y, como una pequeña muestra de lo que se puede hacer en este campo, mostramos (figuras 4.16 y 4.17) las cuencas de atracción asociadas al método de Newton aplicado a las funciones $f(z) = \operatorname{sen} z$ y $f(z) = \exp z - 1$.

En el caso de la función trigonométrica la función de iteración es $N_f(z) = z - \operatorname{tg} z$. La ecuación $\operatorname{sen} z = 0$ tiene infinitas soluciones de la forma $z = k\pi i$, $k \in \mathbb{Z}$. Las regiones pintadas con el mismo color en la figura 4.16 están formadas por puntos de partida para los cuales el método de Newton converge a la misma raíz de la ecuación $\operatorname{sen} z = 0$. En concreto, los

puntos en amarillo convergen a $-\pi$, los puntos en cian convergen a 0 y los puntos en magenta convergen a π. Las regiones en negro están formadas o bien por puntos para los cuales el método de Newton no converge a ninguna raíz o bien por puntos que convergen a cualquiera de las otras tres raíces consideradas.

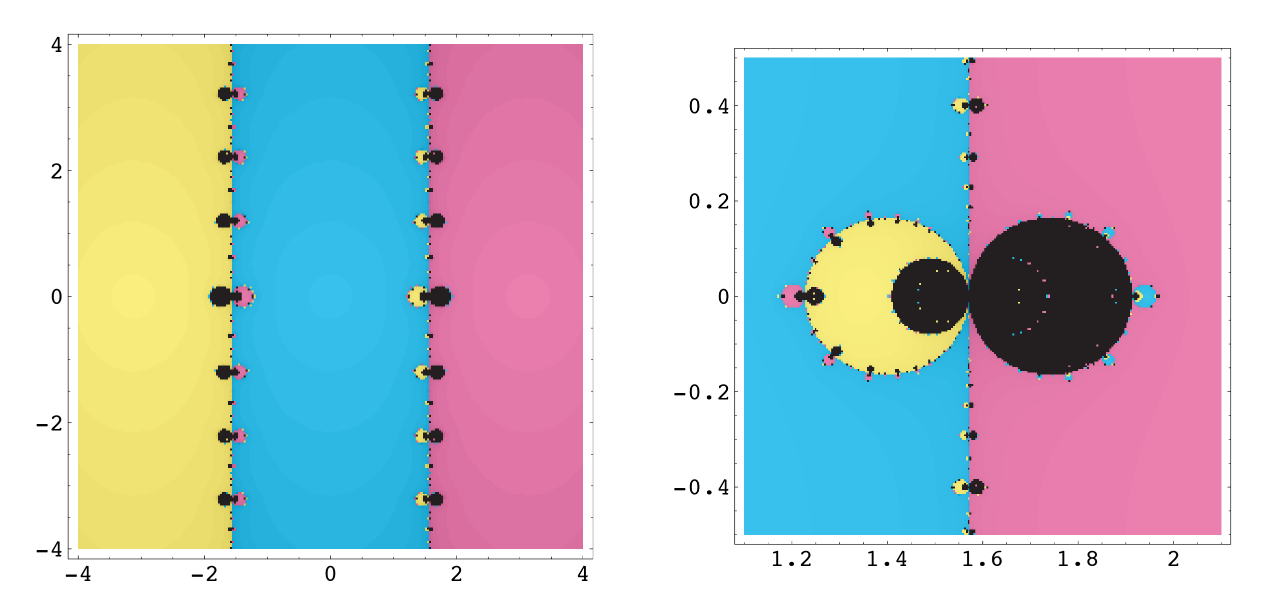


Figura 4.16: A la izquierda se muestran las cuencas de atracción de la función $N_f(z) = z - \operatorname{tg} z$ y a la derecha un detalle de dicha figura.

Notemos que el caso del método de Newton aplicado a la función $f(z) = \cos z$ se puede reducir al caso anterior. Al igual que los ceros de la función $\cos z$ se obtienen aplicando un desplazamiento de $\pi/2$ a los ceros de la función $\operatorname{sen} z$, las cuencas de atracción del método de Newton aplicado a la función $f(z) = \cos z$ son un desplazamiento de $\pi/2$ respecto a las del método de Newton aplicado a la función $f(z) = \operatorname{sen} z$. Así, si denotamos $s(z)$ y $c(z)$ a las funciones de iteración del método de Newton aplicado a las funciones $f(z) = \operatorname{sen} z$ y $f(z) = \cos z$ respectivamente, tenemos que $s(z) = z - \operatorname{tg} z$ y $c(z) = z - \cot z$. Entonces,

$$c(z + \pi/2) = s(z) + \pi/2.$$

En el caso de la función exponencial $f(z) = e^z - 1$, nos encontramos también con infinitas raíces de la forma $z = 2k\pi i$. En detalle de las cuencas de atracción del método de Newton aplicado a esta función, $N_f(z) = z - 1 + e^{-z}$, puede verse en la figura 4.17. Notemos que aparece una gran zona en negro formada por puntos de partida para los cuales no se produce la convergencia ninguna de las raíces. La aparición de tales zonas negras puede estar provocada por el carácter de punto fijo indiferente que tiene el punto del infinito en este caso.

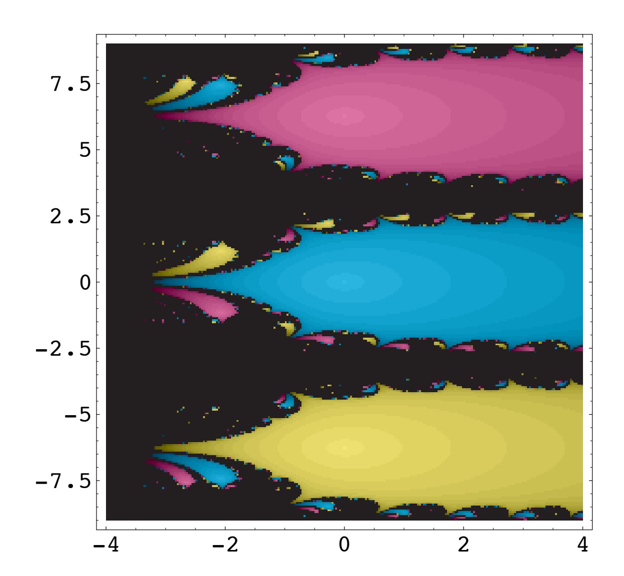


Figura 4.17: Cuencas de atracción del método de Newton aplicado a la función exponencial $f(z) = e^z - 1$.

Capítulo 5

Conjuntos de Julia para polinomios, conjunto de Mandelbrot y método de Newton

5.1. Resultados generales sobre iteración de polinomios

En otras secciones de este texto nos hemos dedicado al estudio de las iteraciones de funciones racionales, tanto en el campo real como en el complejo. Queremos terminar este libro comentando algunos aspectos del comportamiento dinámico de un tipo concreto de funciones racionales, como son las funciones polinomiales. Sin duda, el caso de la iteración de polinomios ha sido el más ampliamente estudiado (véase [16], [26], [49], [73] o [103], por ejemplo). En particular, la iteración de la familia cuadrática $p_c(z) = z^2 + c$ con $c \in \mathbb{C}$, que da lugar al famoso conjunto de Mandelbrot, puede considerarse como la referencia básica de este tipo de problemas. Una introducción asequible al conjunto de Mandelbrot u otras estructuras fractales puede verse en [53], [60] o [93].

En esta sección veremos algunos resultados correspondientes a la dinámica de las iteraciones de esta clase particular de funciones racionales. Recordemos que estamos trabajando en el plano complejo extendido $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Para las iteraciones de polinomios, el punto $z = \infty$ tiene un carácter especial. Sea

$$p(z) = a_d z^d + a_{d-1} z^{d-1} + \cdots + a_1 z + a_0, \quad a_d \neq 0,$$

un polinomio de grado $d \geq 2$.

Denotemos por $C_1 = C_1(p)$ el conjunto de puntos críticos del polinomio p. Como veremos más adelante, este conjunto determina la topología de $\mathcal{J}(p)$.

Teorema 5.1. *Sea p un polinomio de grado $d \geq 2$, entonces*

1) $z = \infty$ es un punto fijo atractor;

2) $z = \infty$ *es un punto crítico de orden* $d - 1$*;*

3) $\mathcal{B}(\infty)^* = \mathcal{B}(\infty)$*, donde* $\mathcal{B}(\infty)^*$ *denota la componente conexa de* $\mathcal{B}(\infty)$ *que contiene a* ∞*, es decir, su cuenca de atracción inmediata.*

El hecho que $z = \infty$ es un punto fijo atractor de un polinomio, implica la siguiente definición.

Definición 5.1. *Se define el* conjunto de Julia lleno *de* p *como el conjunto*

$$\mathcal{K}(p) = \{z \in \mathbb{C} : \{p^n(z) : n \in \mathbb{N}\} \text{ está acotado}\} = \{z \in \mathbb{C} : p^n(z) \nrightarrow \infty\}. \tag{5.1}$$

En otras palabras, $\mathcal{K}(p) = \overline{\mathbb{C}} - \mathcal{B}(\infty)$*, y su* conjunto de Julia *es*

$$\mathcal{J}(p) = \partial\mathcal{K}(p) = \partial(\mathbb{C} - \mathcal{B}(\infty)). \tag{5.2}$$

Es claro que $\mathcal{B}(\infty)$ y $\mathcal{K}(p)$ son invariantes, y en consecuencia $\mathcal{J}(p)$ también lo es.

Corolario 5.2. *El conjunto de Julia* $\mathcal{J}(p)$ *de* p *está acotado. En particular,* $\mathcal{J}(p)$ *no tiene puntos interiores.*

Teorema 5.3. *El conjunto de Julia* $\mathcal{J}(p)$ *es conexo si y sólo si* $\mathcal{B}(\infty) \cap C_1 = \emptyset$*. Equivalentemente, todos los puntos críticos tienen sus órbitas acotadas.*

Corolario 5.4. *Si* $C_1 \subset \mathcal{K}(p)$*, entonces* $\mathcal{J}(p)$ *es conexo, y* $\mathcal{B}(\infty)$ *es simplemente conexo.*

Recordemos que un conjunto $B \subset \mathbb{C}$ tiene medida de Lebesgue nula, y escribimos $med(B) = 0$ si, para cada $\varepsilon > 0$ dado, es posible obtener una sucesión de cuadrados abiertos (cerrados) $Q_1, Q_2, \ldots, Q_i, \ldots$, tales que

$$B \subset \bigcup_{i=1}^{\infty} Q_i \quad \text{y} \quad \sum_{i=1}^{\infty} \text{área}(Q_i) < \varepsilon.$$

El siguiente teorema nos dice que bajo ciertas condiciones el conjunto de Julia, es computacionalmente despreciable.

Teorema 5.5. *Si* $p(z)$ *es un polinomio tal que* $C_1 \subset \mathcal{B}(\infty)$*, entonces*

(a) $\mathcal{J}(p)$ *es totalmente disconexo;*

(b) $\mathcal{J}(p)$ *tiene medida de Lebesgue nula en* $\mathbb{C}$*.*

La condición del teorema nos dice que los iterados de los puntos críticos de p deben converger a ∞. Esto es fácil de detectar computacionalmente, basta calcular los módulos de las sucesivas iteraciones de los puntos críticos y ver que ellos van creciendo en forma indefinida.

5.2. Conjunto de Mandelbrot

El caso de polinomios cuadráticos es particularmente interesante por la propiedad de dicotomía que posee el conjunto de Julia.

Dado un polinomio cuadrático $p(z) = Az^2 + Bz + C$, este tiene un único punto crítico, c_p. Por lo tanto, por el teorema 5.3 se tiene que

Teorema 5.6. *$\mathcal{J}(p)$ es conexo si y sólo si $c_p \in \mathcal{K}(p)$. Además, si $c_p \notin \mathcal{K}_c$, entonces $\mathcal{J}(p)$ es totalmente disconexo, de hecho un conjunto de Cantor.*

Por otra parte, vimos que existe una conjugación afín $\varphi(z) = \alpha z + \beta$ de modo que $\varphi \circ p \circ \varphi^{-1} = p_c$, donde $p_c(z) = z^2 + c$, para algún $c \in \mathbb{C}$. Por lo tanto, para estudiar la dinámica de los polinomios cuadráticos es suficiente estudiar la dinámica de los polinomios p_c, con $c \in \mathbb{C}$.

Usando la dicotomía dada en el teorema 5.6, y el hecho que cada polinomio cuadrático tiene su dinámica representada por algún elemento de la familia cuadrática a paramétrica $p_c(z) = z^2 + c$, se define el conjunto de Mandelbrot de p_c como

$$\mathcal{M} = \{c \in \mathbb{C} : \mathcal{J}(p_c) \text{ es conexo}\}. \tag{5.3}$$

Esta es una definición basada en la topología del conjunto de Julia $\mathcal{J}(p_c)$, el cual denotamos simplemente por $\mathcal{J}_c$. Para que esta definición tenga un carácter más amigable, veamos bajo que condiciones el único punto crítico (finito) $z = 0$ del p_c pertenece al conjunto de Julia lleno $\mathcal{K}_c$.

Tenemos que los puntos críticos de p_c son dados por la ecuación $p_c'(z) = 0$, ellos son $z = 0$ y $z = \infty$. Ahora, los iterados de $z =$ por p_c son dados por $c_0 = p_c(0) = c$, $c_1 = p_c^2(0) = p_c(p_c(0)) = c^2 + c$, $c_3 = p_c^3(0) = p_c(p_c^2(0)) = (c^2 + c)^2 + c, \ldots$, luego $0 \in \mathcal{K}_c$ si y sólo si todos sus iterados permenecen acotados, por lo tanto, tenemos la siguiente caracterización del conjunto de Mandelbrot

$$\begin{aligned} \mathcal{M} &= \{c \in \mathbb{C} : \text{la sucesión } \{c_n\}_{n\in\mathbb{N}} = \{p_c^n(0)\}_{n\in\mathbb{N}} \text{ permanece acotada}\} \\ &= \{c \in \mathbb{C} : c_n = p_c^n(0) \nrightarrow \infty\}. \end{aligned}$$

Sobre el conjunto de Mandelbrot se tiene bastante información, por ejemplo, $\mathcal{M}$ es conexo, compacto y $\mathbb{C} - \mathcal{M}$ es conexo (en otras palabras es un conjunto lleno) [24]. Este conjunto, al igual que los de Julia son, en general, un fractal. La generación de la imagen computacional del conjunto de Maldelbrot es fácil y está basada en su propia definición, usando el algorítmo de escape al infinito.

El plano de los parámetros $c \in \mathbb{C}$ es donde «vive» el conjunto de Mandelbrot, el cual es de hecho un catálogo de la dinámica de polinomios cuadráticos, a cada valor de $c \in \mathbb{C}$ le corresponde un polinomio $p_c(z) = z^2 + c$ cuya dinámica deseamos comprender. Por otra

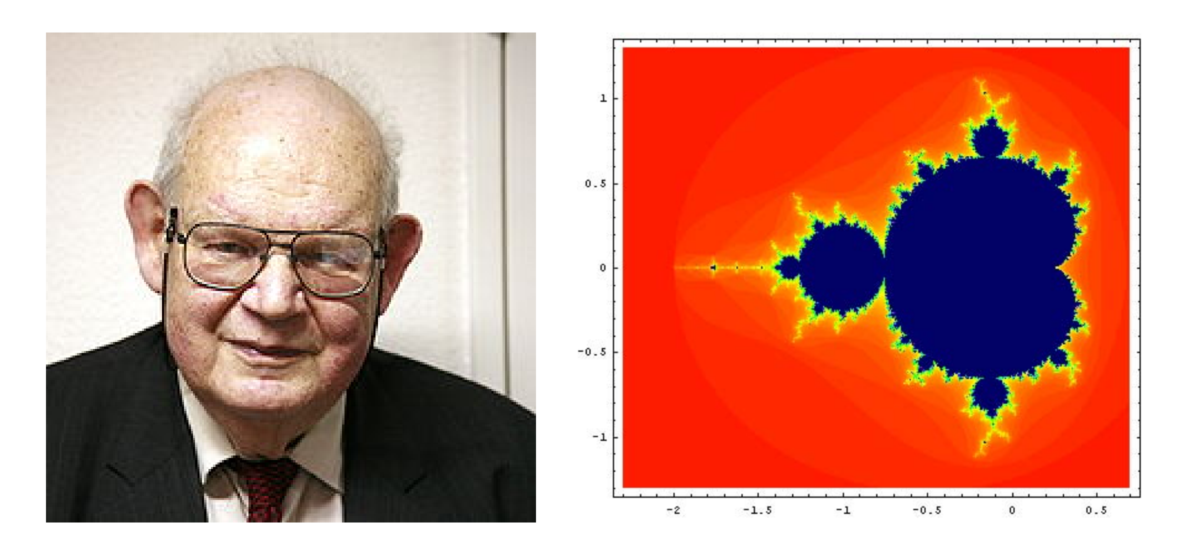


Figura 5.1: Benoît Mandelbrot (1924–2010) y el conjunto que lleva su nombre.

parte, la frontera del conjunto de Mandelbrot corresponde al conjunto de parámetros para los cuales ocurre una bifurcación, esto significa un cambio cualitativo de la dinámica del polionmio, o en otras palabras, si Λ es una componente conexa del interior del conjunto de Mandelbrot y $c_1, c_2 \in \Lambda$, entonces p_{c_1} y p_{c_2} son cuasiconformemente equivalentes. Esto significa que existe un homeomorfismo cuasiconforme $\varphi : \mathbb{C} \to \mathbb{C}$ tal que $\varphi \circ p_{c_1} = p_{c_2} \circ \varphi$. Claro está nos falta definir cuando un homeomorfismo es *cuasiconforme*. Para simplificar, suponemos que $\varphi : U \to V$, donde $V, U \subset \mathbb{C}$ son dominios abiertos es un difeomorfismo (derivable con inversa derivable). La aplicación $\mathbb{R}$–lineal $D\varphi(z) : \mathbb{C} \to \mathbb{C}$ lleva círculos en elipses, sea entonces $E_{\varphi(z)} = D\varphi(z)(S^1)$, definimos el *coeficiente de dilatación* de φ en z, el cual denotamos por $\mathcal{D}_\varphi(z)$, como el cuociente entre la longitud del eje mayor y la longitud del eje menor $E_{\varphi(z)}$, y decimos que φ es cuasiconforme si existe una constante $K \geq 0$ tal que $\mathcal{D}_\varphi = \sup\{\mathcal{D}_{\varphi(z)} : z \in U\} \leq K$. Para la definición más detallada de funciones cuasiconformes ver [45].

5.3. Método de Newton y conjuntos de Julia

Parece ser algo soprendente que los conjuntos de Julia llenos de polinomios cuadráticos aparecen como parte del conjunto de Fatou del método de Newton aplicado a polinomios cúbicos. Esto tiene su asidero en la teoría de «funciones de tipo polinómico», traducción adaptada del nombre en inglés «polynomial-like mappings», desarrollada por A. Douady y J. Hubbard en [45].

Definición 5.2. *Una* función de tipo polinómico *de grado $d \geq 2$ es una terna (f, U, V), donde U y V son conjuntos abiertos de $\mathbb{C}$ isomorfos a discos con $\overline{U} \subset V$ y $f : U \to V$ es una aplicación holomorfa, tal que cada punto en V tiene exactamente d preimágenes en U, cuando ellas son contadas con sus respectivas multiplicidades.*

Ejemplos obvios de tales funciones son los polinomios de grado $d \geq 2$ restringidos a un conjunto abierto suficientemente grande.

Cabe notar que una función de tipo polinómico de grado 2 necesariamente tiene un único punto crítico en el interior de U.

Existen ejemplos de polinomios cúbicos los cuales restringidos a un conjunto abierto son funciones de tipo polinómico de grado 2. Para ello basta considerar un polinomio cúbico $p(z)$ con un punto crítico w_1 escapando al infinito por iteraciones bajo p mientras que las iteraciones del otro punto crítico w_2 permanecen acotadas.

El *conjunto de Julia lleno* para una función de tipo polinómico se define de modo análogo al conjunto de Julia lleno de un polinomio. Esto es, si $f : V \to U$ es una función de tipo polinómico, el conjunto de Julia lleno de f es

$$\mathcal{K}(f) = \{z \in U : f^n(z) \in U \text{ para todo } n \geq 1\}$$

o equivalentemente

$$\mathcal{K}(f) = \bigcap_{n \geq 0} f^{-n}(V)$$

y su conjunto de Julia como

$$\mathcal{J}(f) = \partial\mathcal{K}(f).$$

Sean $f, g : \mathbb{C} \to \mathbb{C}$, si en la ecuación de conjugación $h \circ f = g \circ h$, donde $h : \mathbb{C} \to \mathbb{C}$ es ahora una función cuasiconforme, decimos que f y g son *cuasiconformemente conjugadas*

El siguiente resultado nos muestra que efectivamente, una función de tipo polinómico se comporta como un polinomio.

Teorema 5.7 (Douady y Hubbard, [45])**.** *Sea (f, U, V) una función de tipo polinómico de grado 2. Entonces existe $c \in \mathbb{C}$, tal que f y p_c son cuasiconformemente conjugadas en entornos de sus respectivos conjuntos de Julia llenos. Además, si $\mathcal{K}(f)$ es conexo, entonces tal c es único.*

En particular, se deduce de este teorema que los conjuntos de Julia llenos $\mathcal{K}(p_c)$ y $\mathcal{K}(f)$ son homeomorfos, y más aún, son cuasiconformemente homeomorfos. Esto explica porqué en muchos casos aparecen copias (ligeramente deformadas) de conjuntos de Julia llenos de funciones cuadráticas en la dinámica de otras funciones. Por ejemplo, en la figura 5.2 vemos aparecer el conocido como «conejo de Douady» en el conjunto de Julia del método de Newton para el polinomio

$$p(z) = z^3 - (\lambda - 1)z - \lambda, \quad \lambda = 0.3597736432 + 0.003366617600i. \tag{5.4}$$

Si consideramos ahora la familia de polinomios

$$p_a(z) = z^3 + (a - 2)z^2 - az + a, \tag{5.5}$$

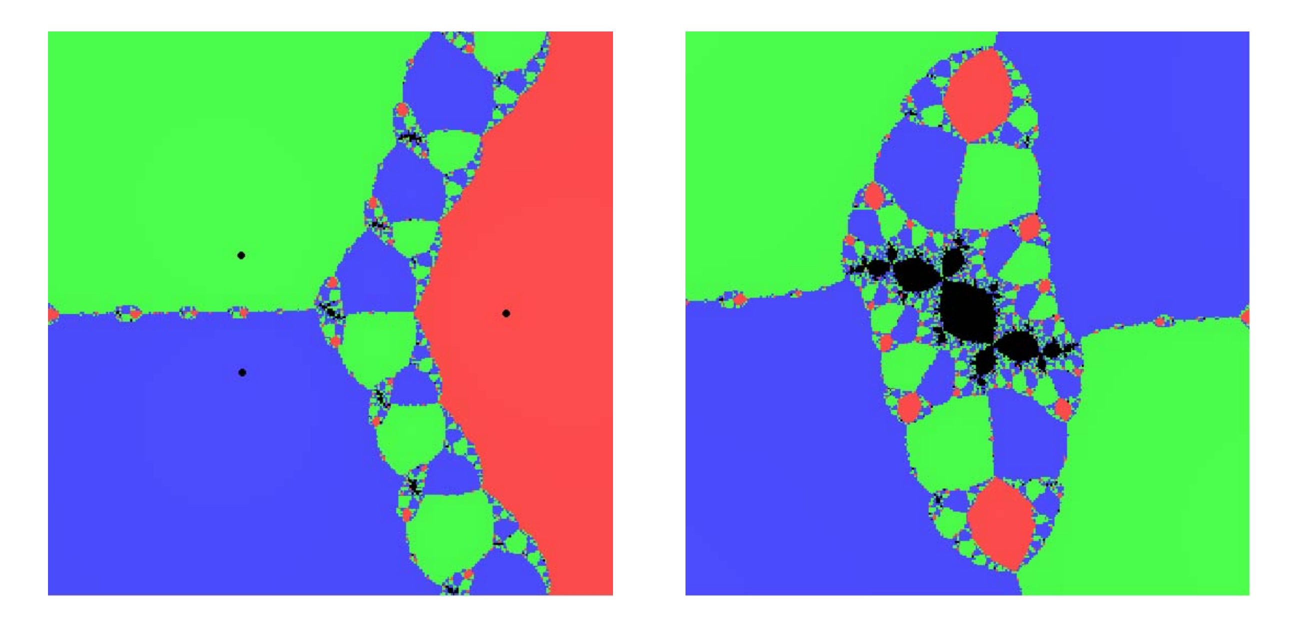


Figura 5.2: El «conejo de Douady» en el conjunto de Julia del método de Newton para el polinomio definido en (5.4). La figura de la derecha muestra una ampliación cerca del origen.

donde

$$a = 1/2(1 + \sqrt{1 + 32/(4 - \ell)}), \quad \ell = \frac{1}{2} + \frac{\sqrt{3}}{2}i,$$

vemos aparecer los conjuntos que se muestran en la figura 5.3. A la derecha se representan las cuencas de atracción de las raíces y su conjunto de Julia, a la izquierda una ampliación de la región central en negro.

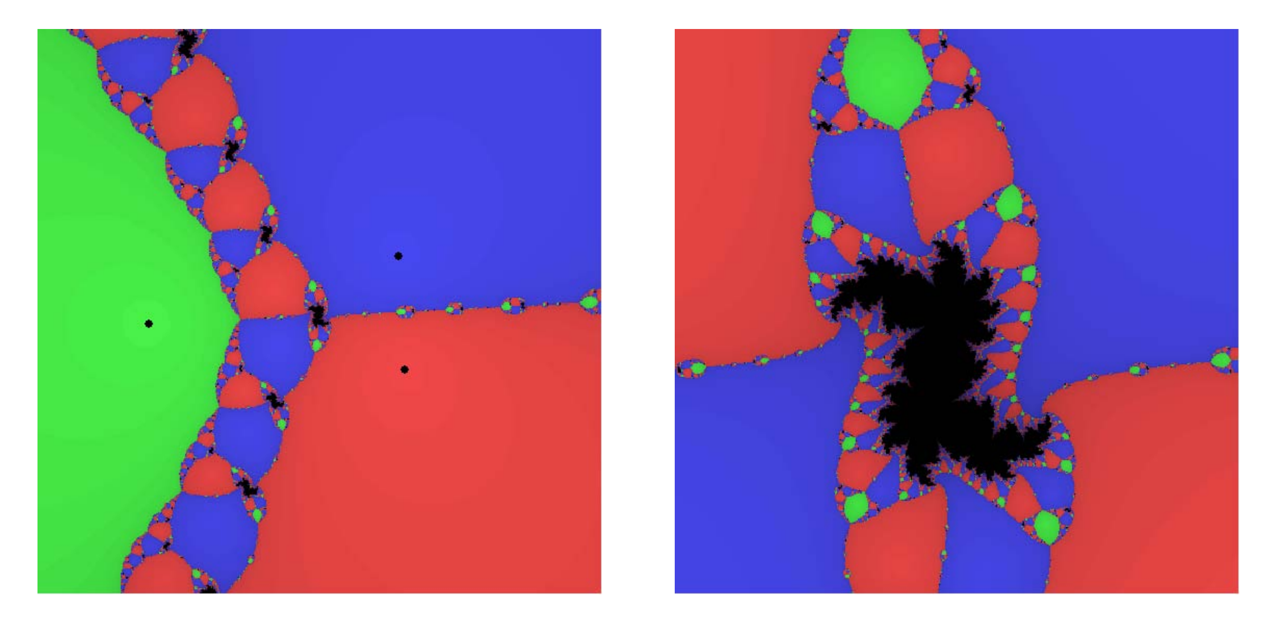


Figura 5.3: Cuencas de atracción de la función de iteración del método de Newton, N_{p_a} donde p_a está definido en (5.5) y una ampliación de la zona central.

Como hemos visto, la teoría de funciones de tipo polinómico explica el hecho que los conjuntos de Julia de polinomios aparecen en los conjuntos de Julia de otras funciones. Esto

se extiende a funciones trascendentales, como puede verse en las figuras 5.4 y 5.5, donde se muestran los conjuntos de Julia asociados a la iteración de funciones trigonométricas (véase también [87] y [137]).

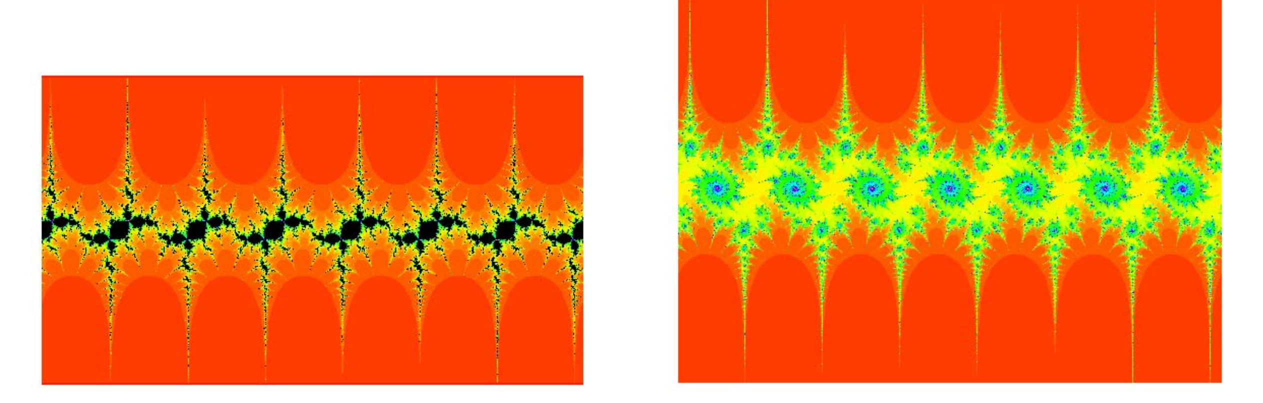


Figura 5.4: Conjunto de Julia de $F(z) = \alpha_1 \operatorname{sen} z$ y de $F(z) = \alpha_2 \operatorname{sen} z$ donde $\alpha_1 = 1.88853 + 0673125i$ y $\alpha_2 = 1 + 0.5i$.

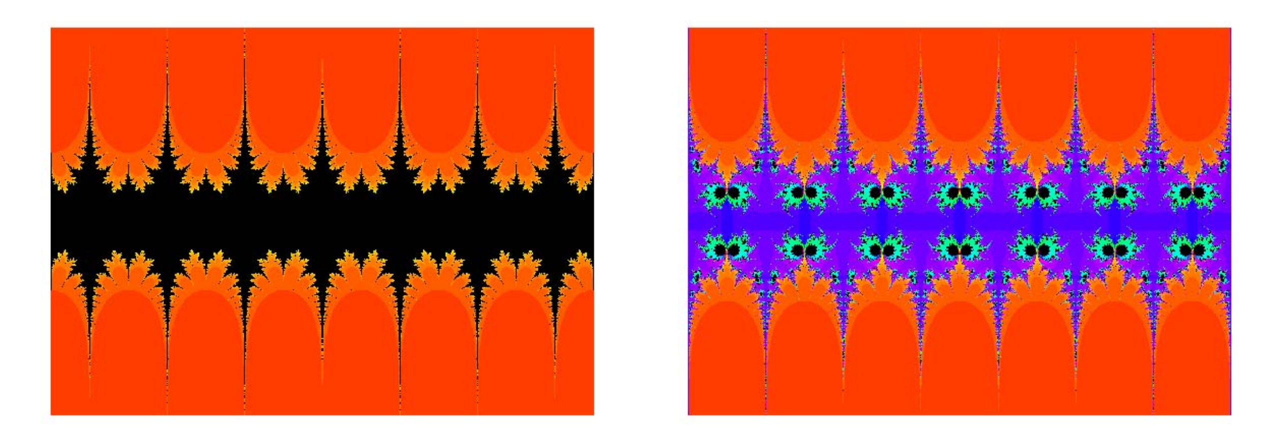


Figura 5.5: Conjunto de Julia de $F(z) = \cos z$ y de $F(z) = 0.67i \cos z$.

5.4. El fractal de Chicho

No nos resistimos a terminar estas notas sin hacer referencia a un problema especialmente querido para los autores. Se trata del conocido como *fractal de Chicho* y está estrechamente vinculado con el contenido de esta sección, ya que puede verse como una modificación del conjunto de Mandelbrot. Su representación gráfica y algunas de sus propiedades matemáticas fueron analizadas en [19], una publicación contenida en un libro publicado en homenaje a José Javier (Chicho) Guadalupe (1945–2000).

La idea consiste en modificar la función de iteración que da lugar al conjunto de Mandelbrot de la siguiente manera. Para $p_c(z) = z^2 + c$, con $z = x + iy$ y $c = a + ib$, podemos

escribir la sucesión de puntos $\{p_c^n(0)\}_{n\in\mathbb{N}}$ en forma recursiva de la siguiente manera:

$$x_{n+1} = x_n^2 - y_n^2 + a, \qquad y_{n+1} = 2x_n y_n + b, \qquad x_0 = y_0 = 0. \tag{5.6}$$

La siguiente variante implícita de la sucesión (5.6) fue propuesta por el propio Chicho y M. Benito en [18]:

$$x_{n+1} = x_n^2 - y_n^2 + a, \qquad y_{n+1} = 2x_{n+1} y_n + b, \qquad x_0 = y_0 = 0. \tag{5.7}$$

La sucesión (5.7) se corresponde con la órbita del punto $(0,0)$ mediante la función de iteración

$$T_c(x,y) = \left(x^2 - y^2 + a, 2(x^2 - y^2 + a)y + b\right), \quad c = a + ib. \tag{5.8}$$

En concreto definimos el *conjunto de Chicho* como el subconjunto de $\mathbb{C}$ formado por los puntos c para los cuales la órbita de $(0,0)$ por T_c está acotada. Denotamos a dicho conjunto como $\mathcal{CH}$. Su representación gráfica puede verse en la figura 5.6.



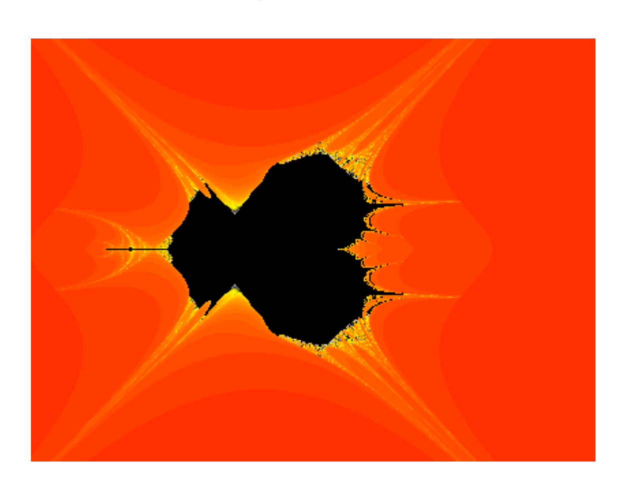


Figura 5.6: Chicho Guadalupe y el fractal que lleva su nombre.

Notemos que el conjunto de Chicho $\mathcal{CH}$ puede interpretarse como la aplicación de un proceso de Gauss-Seidel a la función $p_c(z) = z^2 + c$ de la que resulta el conocido conjunto de Mandelbrot. Las consecuencias que desde el punto de vista dinámico tiene este proceso de «gausseidelización» no son en absoluto triviales.

Así por ejemplo, es bien conocido [16] que el conjunto de Mandelbrot está contenido en el disco de centro el origen y radio 2. Sin embargo, no se ha podido probar o refutar si el conjunto de Chicho está acotado o no. Algunos resultados parciales a este respecto pueden verse en [19]. En concreto, se tiene que:

1. Cuando $c \in \mathbb{R}$, la sucesión $\{u_n\}_{n\in\mathbb{N}} = \{T_c^n(0)\}_{n\in\mathbb{N}}$ es una sucesión de números reales que resulta ser la misma que la definida por p_c para $c \in \mathbb{R}$. Por tanto, $(u_n)_{n\in\mathbb{N}}$ está acotada si y sólo si
$$-2 \le c \le \frac{1}{4}.$$

2. El conjunto de puntos $c = a + bi \in \mathbb{C}$ para los cuales la función T_c definida en (5.8) tiene un punto fijo atractor se sitúa en el interior de la curva

$$\begin{aligned}(r^2-2)r\cos 2\theta - 2r - r^3 + 8\cos\theta \\ = r|\operatorname{sen}\theta|\sqrt{2(8+4r^2+r^4-16r\cos\theta - r^2(r^2-4)\cos 2\theta)},\end{aligned} \tag{5.9}$$

donde $a = (1 - r^2\cos 2\theta)/4$ y $b = (r^2 \operatorname{sen} 2\theta)/4$. La gráfica de la función definida de forma implícita en (5.9) es la curva en forma de corazón que se muestra en la figura 5.7. El conjunto de puntos $c = a+ib$ para los cuales la función T_c tiene un punto fijo atractor se sitúa en el interior de dicho corazón.

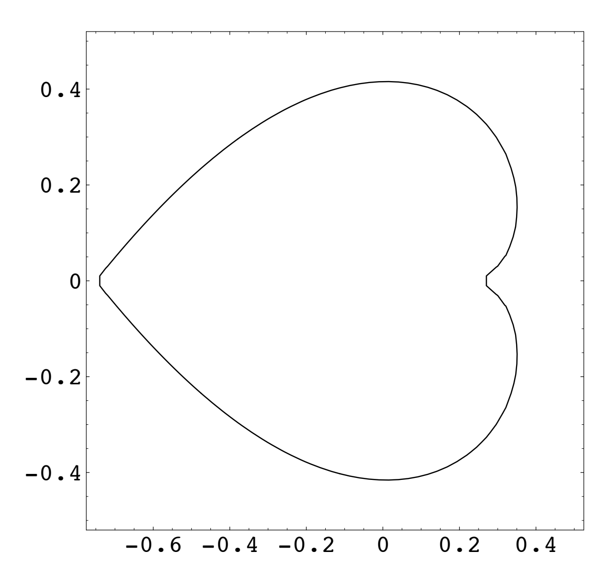


Figura 5.7: Región donde T_c definida en (5.8) tiene un punto fijo atractor.

A continuación podríamos proceder a estudiar los puntos fijos de T_c^2 para caracterizar las órbitas que convergen a un 2-ciclo y así sucesivamente. Sin embargo, esto resulta extremadamente complejo, por lo que nos conformamos con contemplar el aspecto que presenta el conjunto de Chicho que se muestra en la figura 5.6, donde están representados todos aquellos puntos para los cuales la órbita de 0 por T_c se encuentra en el disco de centro el origen y radio 2 después de un número prefijado de iteraciones.

El proceso de «gausseidelización» aplicado en el conjunto de Chicho se puede aplicar a otras funciones de iteración, distintas del polinomio $p_c(z) = z^2 + c$, tal y como se hace en [59]. Desde un punto de vista computacional, puede parecer que la «gausseidelización» es una clara mejora en la eficiencia y en la velocidad de convergencia de un proceso, ya que para el cálculo de y_{n+1} se emplea la actualización x_{n+1} en vez de x_n. Sin embargo, esta afirmación necesita de una prueba rigurosa que establezca bajo que condiciones es cierta. Del mismo modo que se sabe que para resolver sistemas de ecuaciones lineales el método iterativo de Gauss-Seidel no siempre es más rápido que el método de Jacobi, habría que determinar bajo

que condiciones la «gausseidelización» de un proceso iterativo da lugar efectivamente a una mejora computacional.

Por último, en [59] también se pone de manifiesto que la «gausseidelización» de un proceso iterativo tiene una clara influencia en la forma de las cuencas de atracción asociadas a los puntos fijos del proceso. En concreto, en dicho artículo se presentan unos experimentos numéricos que sugieren que la «gausseidelización» tiene el efecto de disminuir la dimensión fractal de los conjuntos de Julia asociados a las citadas cuencas de atracción. De una manera coloquial, podríamos decir que dichos conjuntos de Julia tienen una apariencia menos intrincada. Pero, hasta donde nosotros conocemos, éste es un problema abierto que está a la espera de su solución.

Bibliografía

[1] F. Aguiló y A. Miralles: Consideraciones geométricas acerca del método de Newton, *La Gaceta de la RSME* **7** (2004), n.º 1, 247–260.

[2] L. Ahlfords: *Complex Analysis*, MacGraw–Hill, Nueva York, 1979.

[3] K. Alligood, T. Sauer y J. Yorke: *Chaos: an introduction to dynamical systems*, Springer-Verlag, Berlin-Heidelberg, 1997.

[4] S. Amat, S. Busquier y J. M. Gutiérrez: Geometric constructions of iterative functions to solve nonlinear equations, *J. Comput. Appl. Math.* **157** (2003), 197–205.

[5] I. K. Argyros y F. Szidarovszky: *The theory and application of iteration methods*, C.R.C. Press Inc., Boca Raton, Florida, 1993.

[6] I. K. Argyros y J. M. Gutiérrez: A unified approach for enlarging the radius of convergence for Newton's method and applications, *Nonlinear Functional Analysis and Applications* **10** (2005), 555–563.

[7] D. F. Bailey: A Historical Survey of Solution by Functional Iteration, *Math. Magazine* **62** (1989), n.º 3, 155–166.

[8] F. Balibrea, J. O. Freitas y J. Sousa Ramos: Newton maps for quintic polynomials, *arXiv:math.DS/0501327* (2005), 1–17.

[9] S. Banach: *Théorie des opérations linéaires*, Monografie Matematyczne, Varsovia, 1932.

[10] J. Banks, J. Brooks, G. Cairns, G. Davis y P. Stacey: On Devaney's definition of chaos, *Amer. Math. Monthly* **99** (1992), 332–334.

[11] B. Barna: Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen. I, *Publ. Math. Debrecen* **3** (1953), 109–118.

[12] B. Barna: Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischen Gleichungen. II, *Publ. Math. Debrecen* **4** (1956), 384–397.

[13] B. Barna: Über die divergenzpunkte des Newtonschen verfahrens zur bestimmung von wurzeln algebraischer gleichungen. III, *Publ. Math. Debrecen* **8** (1961), 193–207.

[14] B. Barna: Über die divergenzpunkte des Newtonschen verfahrens zur bestimmung von wurzeln algebraischer gleichungen. IV, *Publ. Math. Debrecen* **14** (1967), 91–97.

[15] M. Barnsley: *Fractals everywhere*, Academic Press, Boston, 1988.

[16] A. F. Beardon: *Iteration of rational functions*, Springer-Verlag, Nueva York, 1991.

[17] A. Ben-Israel: Newton's method with modified functions, *Contemporary Mathematics* **204** (1997), 30–50.

[18] M. Benito y J. J. Guadalupe: Dibujando mediante iteraciones, *Números* **42** (2000), 15–28.

[19] M. Benito, J. M. Gutiérrez y V. Lanchares: El fractal de Chicho, *Margarita Mathematica en memoria de José Javier (Chicho) Guadalupe Hernández*, Serv. Publicaciones Univ. La Rioja, Logroño (2001), 247–254.

[20] W. Bergweiler: Iteration of meromorphic functions, *Bull. Amer. Math. Soc.* **29** (1993), n.º 2, 151–188.

[21] P. Blanchard: Complex Analytic Dynamics on the Riemann sphere, *Bull. Amer. Math. Soc.* **11** (1984), n.º 1, 85–141.

[22] P. Blanchard y A. Chiu: *Complex Dynamics: an informal discussion*, Fractal Geometry and Analysis, Eds. J. Bélair & S. Dubuc, Kluwer Academic Publishers (1991), 45–98.

[23] L. E. Böttcher: The principal laws of convergence of iterates and their application to Analysis, *Izv. Kasan. Fiz.-Mat. Obshch* **14** (1904), 155–234.

[24] B. Branner: The Mandelbrot set, *Proc. Symp. Applied Math.* (1989), 75–105.

[25] F. Cajori, Historical note on the Newton-Raphson method of approximation, *Amer. Math. Monthly* **18** (1910), 29–33.

[26] L. Carleson y T. Gamelin: *Complex Dynamics*, Springer-Verlag, Berlín-Heidelberg, 1993.

[27] B. Casselman: YBC 7289, a precursor of the Euclid's Elements of Geometry, `http://www.math.ubc.ca/~cass/Euclid/ybc/ybc.html`

[28] A. Cayley: The Newton-Fourier imaginary problem, *Amer. J. Math.* **2** (1879), 97–97.

[29] A. Cayley: Application of the Newton-Fourier method to an imaginary root of an equation, *Quaterly J. Pure Appl. Math.* **16** (1879), 179–185.

[30] A. Cayley: Sur les racines d'une équation algébrique, *Comptes Rendus Acad. Sci.* **110** (1890), 215–218.

[31] J. L. Chabert et al.:*A History of Algorithms: from the Pebble to the Microchip*, Springer-Verlag, Berlín-Heidelberg, 1999.

[32] S. Chandrasekhar: *Radiative transfer*, Dover, Nueva York, 1960.

[33] E. D. Charles y J. B. Tatum: The convergence of Newton-Raphson iteration with Kepler's equation, *Celestial Mechanics and Dynamical Astronomy* **69** (1998), 357–372.

[34] P. Colwell: *Solving Kepler's equation over three centuries*, Willmann-Bell, Inc., Richmond, VA, 1993.

[35] B. A. Conway: An improved algorithm due to Laguerre for the solution of Kepler's equation, *Celest. Mech.* **39** (1986), 199–211.

[36] M. Cosnard y C. Masse: Convergence presque partout de la méthode de Newton, *C. R. Acad. Sc. Paris* **297** (1983), 549–552.

[37] J. H. Curry, L. Garnett y D. Sullivan: On the iteration of rational functions: Computer experiments with Newton's method, *Commun. Math. Phys.* **91** (1983), 267–277.

[38] J. M. A. Danby y T. M. Burkardt: The solution of Kepler equation I, *Celestial Mechanics* **40** (1983), 95–107.

[39] J. M. A. Danby y T. M. Burkardt: The solution of Kepler equation III, *Celestial Mechanics* **31** (1987), 303–312.

[40] J. P. Dedieu: *Points fixes, zéros et la Méthode de Newton*, Springer-Verlag, Berlín-Heidelberg, 2006.

[41] J. E. Dennis y R. B. Schnabel: *Numerical methods for unconstrained optimization and nonlinear equations*, Classics in Applied Mathematics, Vol. 16, SIAM, Filadelfia, 1996.

[42] R. L. Devaney: *A first course in Chaotic Dynamical Systems*, Addison-Wesley, Redwood City (CA), 1992.

[43] R. L. Devaney: *An Introduction to Chaotic Dynamical Systems, Second Edition*, Westview Press, Cambridge, 2003.

[44] L. E. Dickson: *Modern Algebraic Theories*, H. Sanborn and Co., Chicago, 1926.

[45] A. Douady y J. Hubbard: On the dynamics of polynomial-like mappings, *Ann. Scient. Ec. Norm. Sup. 4e series* **18** (1985), 287–343.

[46] J. A. Ezquerro, J. M. Gutiérrez, M. A. Hernández, N. Romero y M. J. Rubio: El método de Newton: de Newton a Kantorovich, *La Gaceta de la RSME* **13** (2010), n.º 1, 53–76.

[47] J. A. Ezquerro, J. M. Gutiérrez, M. A. Hernández, N. Romero y M. J. Rubio: Relaciones de recurrencia en el método de Newton-Kantorovich, *Contribuciones científicas en honor de Mirian Andrés Gómez*, Serv. Publicaciones Univ. La Rioja, Logroño, (2010), 319–333.

[48] N. Fagella: Invariants en dinàmica complexa, *Bull. Soc. Mat. Cat.* **23** (2007), n.º 1, 29–51.

[49] N. Fagella y X. Jarque: *Iteración compleja y fractales*, *Vicens Vives*, Barcelona, 2007.

[50] J. D. Faires y R. L. Burden: *Métodos Numéricos, 3a Ed.*, Thomson, Madrid, 2004.

[51] P. Fatou: Sur les équations fonctionelles, *Bull. Soc. Math. France* **47** (1919), 161–271.

[52] P. Fatou: Sur les équations fonctionelles, *Bull. Soc. Math. France* **48** (1920), 208–314.

[53] M. Frame y B. B. Mandelbrot: *Fractals, graphics and mathematics education*, Mathematical Association of America, Washington, DC, 2002.

[54] Gallica-Math: OEuvres complètes: Biblioteca numérica Gallica de la Bibliothèque Nationale de France, `http://mathdoc.emath.fr/OEUVRES/`

[55] W. J. Gilbert: The complex dynamics of Newton's method for a double root, *Computers Math. Applic.* **22** (1991), n.º 10, 115–119.

[56] A. Giraldo y M. A. Sastre: *Sistemas dinámicos discretos y caos. Teoría, ejemplos y algoritmos*, Fundación General de la Universidad Politécnica de Madrid, Madrid, 2002.

[57] D. Gulick: *Encounter with chaos*, McGraw Hill, Nueva York, 1992.

[58] J. M. Gutiérrez, M. A. Hernández y M. A. Salanova: Calculus of nth roots and third order iterative methods, *Nonlinear Analysis* **47** (2001), 2875–2880.

[59] J. M. Gutiérrez, Á. A. Magreñán y J. L. Varona: The “Gauss-Seidelization” of iterative methods for solving nonlinear equations in the complex plane, *Appl. Math. Comput.* **218** (2011), 2467–2479.

[60] M. de Guzmán, M. Á. Martín, M. Morán y M. Reyes: *Estructuras fractales y sus aplicaciones*, Editorial Labor, Barcelona, 1993.

[61] J. M. Hawkins: McMullen’s root-finding algorithm for cubic polynomials, *Proc. Amer. Math. Soc.* **130** (2002), n.º 9, 2583–2592.

[62] M. Haruta: Newton’s method on the complex exponential function, *Trans. Amer. Math. Soc.* **351** (1999), 2499–2513.

[63] J. Head: *The combinatorics of Newton’s method for cubic polynomials*, Ph. D. Thesis, Cornell Univ., Ithaca (N. Y.), 1987.

[64] P. Henrici: *Elements of Numerical Analysis*, John Wiley & Sons, Inc., Nueva York, 1964.

[65] M. A. Hernández y M. A. Salanova: *La convexidad en la resolución de ecuaciones no lineales*, Servicio de Publicaciones de la Universidad de La Rioja, 1996.

[66] R. A. Holmgren: *A first course in discrete dynamical systems, second edition*, Springer-Verlag, Berlín-Heidelberg, 1996.

[67] P. Horton: No fooling! Newton’s method can be fooled, *Math. Magazine* **80** (2007), 383–387.

[68] A. S. Householder: *The numerical treatment of a single nonlinear equation*, McGraw-Hill, Nueva York, 1970.

[69] J. H. Hubbard y B. West: *Differential equations: a dynamical systems approach*, Springer-Verlag, Berlín-Heidelberg, 1991.

[70] J. H. Hubbard, D. Schleicher y S. Sutherland: How to find all roots of complex polynomials by Newton’s method, *Inv. Math.* **146** (2001), 1–33.

[71] H. Hurley: Multiple attractors in Newton’s method, *Erg. Theory and Dyn. Systems* **6** (1984), 561–569.

[72] G. Julia: Memoire sur l’iteration des fonctions rationelles, *J. de Math. pures et appliquées* **8:1** (1918), 47–215.

[73] B. Kalantari: *Polynomial root-finding and polyomiography*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2009.

[74] L. V. Kantorovich: On Newton's method for functional equations, *Dokl Akad. Nauk SSSR* **59** (1948), 1237–1240 (en ruso).

[75] L. V. Kantorovich: The majorant principle and Newton's method, *Dokl Akad. Nauk SSSR* **76** (1951), 17–20 (en ruso).

[76] L. V. Kantorovich: *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1964.

[77] L. V. Kantorovich y G. P. Akilov: *Functional Analysis*, Pergamon Press, Oxford, 1982.

[78] C. T. Kelley: *Iterative methods for linear and nonlinear equations*, Frontiers in Applied Mathematics, Vol. 16, SIAM, Filadelfia, 1995.

[79] H. B. Keller: *Numerical solution of two point boundary value problem*, SIAM, Filadelfia, 1976.

[80] D. Kincaid y W. Cheney: *Análisis Numérico. Las matemáticas del cálculo científico*, Addison-Wesley Iberoamericana, Wilmington, 1994.

[81] R. F. King: Improving the Van de Vel root-finding method Algorithm, *Computing* **30** (1983), 373–378.

[82] K. Kneisl: Julia sets for the super-Newton method, Cauchy's method and Halley's method, *Chaos* **11** (2001), n.º 2, 359–370.

[83] R. J. Knill: A Modified Babylonian Algorithm, *Amer. Math. Monthly* **99** (1992), 734–737.

[84] N. Kollerstrom: Thomas Simpson and 'Newton's method of approximation': an enduring myth, *British J. Hist. Science* **25** (1992), 347–354.

[85] G. Königs: Recherches sur les èquationes fontionelles, *Ann. l'Ecole Norm.* **1** (1884), Suplement.

[86] P. Kravanja y A. Haegemans: A modification of Newton's method for analytic mappings having multiple zeros, *Computing* **62** (1999), 129–145.

[87] V. Lanchares y I. L. Pérez-Barrón: The dynamics of Kepler equation, en *Analytic and numerical techniques in orbital dynamics, Monogr. Real Acad. Ci. Exact. Fís.-Quím. Nat. Zaragoza* **22** (2002), 75–82.

[88] P. Lancaster y L. Rodman: *Algebraic Riccati equations*, Oxford University Press, Oxford, 1995.

[89] M. S. LATTÈS: Sur l'iteration de Substitutions Rationelles et Fonctions de Poincaré, Comptes Rendus Acad. Sci. **166** (1918), 26–28.

[90] R. LAUBENBACHER, G. MCGRATH Y D. PENGELLEY: Lagrange and the solution of numerical equations, *Historia Math.* **28** (2001), **3**, 220–231.

[91] T. LI Y J. YORKE: Period three implies chaos, *Amer. Math. Monthly* **82** (1975), 985–993.

[92] F. LUCAS: Sur une application de la Mécanique rationnelle à la théorie des équations, *C. R. Hebdomadaires Séances Acad. Sci.* **89** (1879), 224–226.

[93] B. MANDELBROT: *The fractal geometry of Nature*, W. H. Freeman and Co., Nueva York, 1982.

[94] M. MARTELLI: *Introduction to discrete dynamical systems and chaos*, Wiley-Interscience Publ., Nueva York, 1999.

[95] J. H. MATHEWS: Bibliography for Newton's method, `http://math.fullerton.edu/mathews/n2003/Newton'sMethodBib.html`

[96] MATLAB: Repositorio del entorno de programación Matlab, `http://www.mathworks.es/products/matlab/`

[97] R. MAY: Simple mathematical models with very complicated dynamics, *Nature* **261** (1976), 459–467.

[98] M. MCCLURE: Newton's method for complex polynomials, *Mathematica in Education and Research* **11** (2006), n.º 2, 2–15.

[99] C. MCMULLEN: Families of rational maps and iterative root-finding algorithms, *Annals of Mathematics* **125** (1987), 467–493.

[100] C. MCMULLEN: *Complex dynamics and renormalization*, Annals of Mathematics Studies **135**, Princeton University Press, Princeton, New Jersey, 1994.

[101] J. M. MCNAMEE: A bibliography on roots of polynomials: Newton's method, `http://www1.elsevier.com/homepage/sac/cam/mcnamee/02.htm`

[102] J. MILNOR: *Dynamics in one complex variable: Introductory lectures. Third edition*, Princeton University Press, Princeton, New Jersey, 2006.

[103] S. MOROSAWA, Y. NISHIMURA, M. TANIGUCHI Y T. UEDA: *Holomorphic Dynamics*, Cambridge University Press, Cambridge, 2000.

[104] NETLIB: Software repository at the University of Tennessee, http://www.netlib.org/minpack/

[105] O. NEUGEBAUER Y A. SACHS: *Mathematical cuneiform texts*, American Oriental Society, New Haven, Conn., 1945.

[106] K. NISHIZAWA Y M. FUJIMURA: Families of rational maps and convergence basins of Newton's method, *Proc. Japan Acad.* **68** Ser. A, (1992), 143–147.

[107] J. J. O'CONNOR Y E. F. ROBERTSON: The MacTutor History of Mathematics archive: Nine Chapters on the Mathematical Art, http://www.gap-system.org/~history/HistTopics/Nine_chapters.html

[108] OCTAVE: Repositorio del programa libre de cálculo numérico Octave, http://www.gnu.org/software/octave/

[109] J. M. ORTEGA Y W. C. RHEINBOLDT: *Iterative solution of nonlinear equations in several variables*, Academic Press, Nueva York, 1970.

[110] A. OSTROWSKI: Über die Konvergenz und die Abrundungsfestigkeit des Newtonschen Verfahrens, *Rec. Math.* **2** (1937), 1073–1095.

[111] A. OSTROWSKI: Über einen Fall der Konvergenz des Newtonschen Näherungsverfahrens, *Rec. Math.* **3** (1938), 254–258.

[112] A. OSTROWSKI: *Solution of equations and systems of equations*, Academic Press, Nueva York, 1966.

[113] M. PALACIOS: Kepler equation and accelerated Newton method, *J. Comput. Appl. Math.* **138** (2002), 335–346.

[114] H. O. PEITGEN Y P.H. RICHTER: *The beauty of fractals*, Springer-Verlag, Berlín-Heidelberg, 1986.

[115] H. O. PEITGEN, D. SAUPE Y H. JÜRGENS: *Fractals on the classroom. Vol. I: Introduction to fractals and chaos.*, Springer-Verlag, Nueva York, 1992.

[116] PETSC: Portable, Extensible Toolkit for Scientific Computation, http://www.mcs.anl.gov/petsc/petsc-as/

[117] S. PLAZA: *Fractales y generación computacional de imágenes*, Monografía número 16, Instituto de Matemáticas y Ciencias Afines, IMCA, Perú, 2000.

[118] S. PLAZA Y N. ROMERO: Attracting cycles for the relaxed Newton's method, *J. Comput. Appl. Math.* **235** (2011), 3238–3244.

[119] S. Plaza y V. Vergara: Existence of periodic orbit for Newton method, *Scientia, Series A, Mathematical Sciences* **7** (2001), 31–36.

[120] B. T. Polyak: Newton-Kantorovich method and its global convergence, *J. Math. Sciences* **133** (2006), n.º 4, 1513–1523.

[121] F. A. Potra y V. Pták: *Nondiscrete induction and iterative processes*, Pitman, Londres, 1984.

[122] L. B. Rall: Convergence of Newton process to multiple solutions, *Numer. Math.* **9** (1966), n.º 1, 23–37.

[123] L. B. Rall: *Computational solution of nonlinear operator equations*, Robert E. Krieger Publishing Company, Huntington, Nueva York, 1979.

[124] W. C. Rheinboldt: An adaptice continuation process for solving systems of nonlinear equations, *Polish Acad. Sci. Banach Center Publ.* **3** (1977), 129–142.

[125] G. Roberts y J. Horgan-Kobelski: Newton's versus Halley's method: a dynamical systems approach, *Intern. J. Bifurcation Chaos* **14** (2004), n.º 10, 3459–3475.

[126] R. C. Robinson: *Dynamical systems: Stability, Symbolic Dynamics and Chaos. Second Edition*, CRC Press, Nueva York, 1999.

[127] R. C. Robinson: *An introduction to dynamical systems: continuous and discrete*, Pearson Prentice Hall, New Jersey, 2004.

[128] F. M. Rodríguez-Vásquez: *Desarrollo conceptual de los métodos iterativos en la resolución de ecuaciones no lineales: un enfoque didáctico*, Tesis Doctoral, Univ. Salamanca, 2010.

[129] J. Rückert y D. Schleicher: On the Newton's method for entire functions, *J. London Math. Soc.* **75** (2007), n.º 2, 659–676.

[130] Sage: Repositorio del programa de cálculo matemático Sage, `http://www.sagemath.org/`

[131] D. A. Sánchez: An alternative to the shooting method for a certain class of boundary value problems, *Amer. Math. Monthly* **108** (2001), n.º 6, 552–555.

[132] G. Saunder: *Iteration of rational function of one complex variable and basins of attractive fixed points*, Ph. D. Thesis, Univ. of California, Berkeley, 1984.

[133] F. Scheid: *2000 Solved Problems in Numerical Analysis*, Schaum's Solved Problem Series, McGraw–Hill, Nueva York, 1990.

[134] D. Schleicher: Newton's method as a dynamical system: efficient root finding of polynomial and the Riemann ζ function, *Fields Institute Communications* **53** (2008), 1–12.

[135] E. Schröder: Über unendlich viele Algorithmen zur Auflösung der Gleichungen., *Math. Ann.* **2** (1870), 317–365. (Traducido por G. W. Stewart como *On Infinitely Many Algorithms for Solving Equations* en 1992 (revisado en enero de 1993), disponible vía ftp en `ftp://thales.cs.umd.edu` en el directorio `pub/reports`.)

[136] Scilab: The Scilab Consortium, `http://www.scilab.org/`

[137] W. T. Shaw: *Complex Analysis with Mathematica*, Cambridge University Press, Cambridge, 2006.

[138] M. Shishikura: The connectivity of the Julia set and fixed points, en *"Complex dynamics: families and friends"* (Ed. by D. Schleicher), A. K. Peters (2009), 257–276.

[139] S. Smale: On the efficiency of algorithms of analysis, *Bull. Amer. Math. Soc.* **13** (1985), n.º 4, 87–121.

[140] I. Stewart: *Historia de las Matemáticas en los últimos 10000 años*, Crítica, Barcelona, 2008.

[141] S. Strogatz: *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA, 1994.

[142] D. Sullivan: Quasi conformal homeomorphisms and dynamics. I. Solution of Fatou-Julia problem wandering domains, *Ann. Math.* **122** (1985), n.º 2, 401–418.

[143] The 2000 wiewpoints Group: Proof without words: geometric series, *Mathematics Magazine* **74** (2001), 320.

[144] H. Thunberg: Periodicity versus chaos in one–dimensional dynamics, *SIAM Review* **43** (2000), 3–30.

[145] P. Touhey: Yet another definition of chaos, *Amer. Math. Monthly* **104** (1997), 411–415.

[146] J. F. Traub: *Iterative methods for the solution of equations*, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[147] H. Van de Vel: A method for computing a root of a single nonlinear equation, including its multiplicity, *Computing* **14** (1975), 167–171.

[148] M. Vander Straeten y H. Van de Vel: Multiple root-finding methods, *J. Comput. Appl. Math.* **40** (1992), 105–114.

[149] J. L. Varona: Graphic and numerical comparison between iterative methods, *Math. Intelligencer* **24** (2002), n.º 1, 37–46.

[150] Yale University: The Yale Babylonian Collection, http://www.yale.edu/nelc/babylonian.html

[151] T. Yamamoto: Historical developments in convergence analysis for Newton's and Newton-like methods, *J. Comput. Appl. Math.* **124** (2000), 1–23.

[152] L. Yau y A. Ben-Israel: The Newton and Halley methods for complex roots, *Amer. Math. Monthly* **105** (1998), n.º 9, 806–818.

[153] T. J. Ypma: Historical development of the Newton-Raphson method, *SIAM Review* **37** (1995), n.º 4, 531–551.

[154] J. Walsh: The dynamics of Newton's method for cubic polynomials, *College Mathematics Journal* **26** (1995), n.º 1, 22–28.

[155] E. W. Weisstein: Bring-Jerrard Quintic Form. From MathWorld, A Wolfram Web Resource. http://mathworld.wolfram.com/Bring-JerrardQuinticForm.html

[156] Zuse Institute Berlin: Software repository for Peter Deuflhards Book "Newton Methods for Nonlinear Problems – Affine Invariance and Adaptive Algorithms", http://www.zib.de/Numerik/numsoft/NewtonLib/index.en.html

Índice alfabético



Servicio de Publicaciones
Biblioteca Universitaria
C/ Piscinas, s/n
26006 Logroño (La Rioja)
Teléfono: 941 299 187

http://publicaciones.unirioja.es
www.unirioja.es